CRT-672

# ECONOMETRIC MODELS OF ACCIDENT DISTRIBUTIONS

by

M. Boyer*
G. Dionne**
C. Vanasse***

* Département de sciences économiques, Centre de recherche et développement en économique and Centre de recherche sur les transports, Université de Montréal.

** Département de sciences économiques and Centre de recherche sur les transports, Université de Montréal.

*** Centre de recherche sur les transports, Université de Montréal.

We are grateful to the Régie de l'assurance automobile du Québec (R.A.A.Q.), F.C.A.R. (Québec) and S.S.H.R.C. (Canada) for their financial support.

Centre de recherche sur les transports - Publication #672
Département de sciences économiques - Cahier 9001
Centre de recherche et développement en économique - Cahier 0190

January 1990

90-03003/

Dépôt légal - 1er trimestre 1990

## TABLE OF CONTENTS


| | | Page |
|---|---|---|
| INTRODUCTION | | 1 |
| I. | THE DATA | 4 |
| II. | THE ECONOMETRICS OF CATEGORICAL AND COUNT DATA MODELS | 6 |
| | II.1 Count data models without individual characteristics | 6 |
| | II.2 Categorical models: linear probability, probit, logit and multinomial logit models | 8 |
| | II.3 Count data models with individual characteristics in the regression component | 18 |
| | II.4 Related sampling and estimation issues | 23 |
| | II.5 Comparisons of estimated individual probabilities of accident by the different models | 25 |
| III. | THE DETERMINATION OF AN OPTIMAL PRICING FORMULA FOR AUTOMOBILE INSURANCE | 25 |
| CONCLUSION | | 29 |

## RÉSUMÉ

Cet article traite de l'économétrie des accidents d'automobiles, soit de l'estimation des facteurs significatifs qui expliquent les nombres d'accidents sur une base individuelle pour une période donnée. Les nombres d'accidents d'automobiles sont des variables discrètes et, par conséquent, représentent une procédure de comptage : la variable dépendante peut prendre seulement des valeurs discrètes non négatives. Nous utilisons une banque de données comprenant des accidents et des caractéristiques individuelles. La variable dépendante est le nombre d'accidents d'un individu à une période donnée. Les caractéristiques individuelles sont considérées comme exogènes ou prédéterminées. Nous avons estimé quatre modèles avec variables qualitatives (linéaire, probit, logit et logit polytomique) et quatre modèles avec variables discrètes (Poisson et binômiale négative avec et sans composante de régression). Il est difficile de comparer les résultats économétriques des différents modèles étant donné que certains ne sont pas emboîtés. Par contre, il est montré que le modèle binômiale négative avec composante de régression donne une bonne approximation de la vraie distribution d'accidents. Différents tests statistiques rejettent les modèles Poisson (avec ou sans composante de régression) et le modèle binômiale négative sans caractéristiques individuelles. Il est aussi observé que tous les modèles estimés donnent les mêmes résultats qualitatifs (essentiellement les mêmes variables significatives), mais différents lorsque des prédictions de probabilités d'accidents ou de nombres moyens d'accidents sont effectuées. Pour des analyses quantitatives il est important de sélectionner les modèles appropriés. De plus, il est montré que l'expérience de conduite d'un individu est un bon indicateur de risque. Finalement, nous appliquons les résultats statistiques à un modèle de tarification de l'assurance en présence de risque moral.



Mots clés: accidents d'automobiles, modèles à données discrètes, expérience passée, caractéristiques individuelles, tarification de l'assurance, risque moral.



## ABSTRACT

This paper deals with the econometrics of car accidents, that is the estimation of the relative importance or significance of the factors explaining the number of accidents in a given period of time on an individual basis. The number of car accidents is a discrete variable and therefore represents a count process : the dependent variable takes only non-negative integer values. We exploit a data base consisting of the accident record and characteristics of individual drivers. Hence, the observed dependent variable is the number of accidents an individual i had in the time period considered. The individual characteristics are considered exogenous or predetermined and may or may not be significant factors in explaining the number of accidents. We have estimated four categorical models (linear probability, probit, logit and multinomial logit) and four count data models (Poisson and negative binomial models with and without individual characteristics in the regression component). It is difficult to compare the econometric results of the different models since some of these models are not nested. However, it is shown that the negative binomial model with a regression component produces a reasonable approximation of the true distribution of accidents. Different statistical tests reject the Poisson models (with and without a regression component) and the negative binomial model without individual characteristics. It is also observed that all estimated models provide the same qualitative results (essentially the same significant variables), but differ when predictions of either the probabilities of accidents or the expected number of accidents were made. For quantitative predictions, it is important to select the appropriate model. Moreover, it is shown that, in all models, the individual's past driving experience is a good predictor of risk. Finally, we apply the statistical results to a model of insurance rating in the presence of moral hazard.



Key words: Car Accidents, Count Data Models, Accident Record, Individual Characteristics, Insurance Rating, Moral Hazard.

## INTRODUCTION

This paper deals with the econometrics of car accidents, that is the estimation of the relative importance or significance of the factors explaining the number of accidents in a given period of time on an individual basis. The number of car accidents is a discrete variable and therefore represents a count process: the dependent variable takes only non–negative integer values. We exploit a data base consisting of the accident record and characteristics of individual drivers. Hence, the observed dependent variable is the number of accidents an individual i had in the time period considered. The individual characteristics are considered exogenous or predetermined and may or not be significant factors in explaining the number of accidents.

Count variables appear in the economic theory of individual behavior in two different but related sets of situations. In the first set, the individual is directly choosing a naturally discrete variable by selecting one among mutually exclusive possibilities. Examples of the first set of situations are the mutually exclusive choices of a particular brand or of a particular good in a discrete set of available goods, of a particular job among all the ex ante discrete available job opportunities and of a mode of consumption like a transport mode amongst all the ex ante discrete available modes.

In the second set of situations giving rise to count variables, the individual really chooses a continuous variable, but this choice is private information to the agent and as such is unobservable by an outsider. However, the outsider can observe a proxy integer variable (a count variable) of the individual's choice. Examples of the second set of situations are the level of work effort chosen by a worker or manager while an outside observer can only see if the firm succeeded or not in obtaining a contract, in entering a market, in avoiding bankruptcy, in avoiding a strike. Other examples are the probability of illness or health chosen by an insuree (through a proper diet or exercise), while the outside observer or the insurer can only see, that is can only gather data on whether the individual was sick or not, on how many visits he made to a doctor, or on how many drug types he regularly consumes; the probability of divorce chosen by an individual (through a proper level of divorce preventing activities or behavior) while the outside observer can only observe whether the individual had a divorce or not; and so on.

Count data constitute a particularly important form of empirical observations not only in economics but also in all social sciences. Although their systematic use in econometrics is

rather recent, they do have a long and successful history in biological research for which many of the statistical models were developed originally [Patil (1970)]. In economics, count data models have been or could be profitably used to explain the number of product varieties produced by a firm [Holland (1973), for the case of chemical products]; to explain the use of health services such as the number of visits to a doctor, to hospitals, the number of drugs consumed [Cameron, Trivedi, Milne and Piggott (1988)]; the number of units bought of a good [Gilbert (1979)]; the number of patents applied for by firms [Haussman, Hall and Griliches (1984)]; the number of airline incidents [Rose (1988)]; and many other cases relating to the number of strikes, the number of children in a family, the number of new firm created, the number of banks or other firms failures, and the number of layoffs spells. One can think of quite a few other examples; there is no doubt many other applications of count data models could be made in applied econometrics and quantitative social sciences in general.

Although economic theory does provide many such examples of count data, it is in general silent regarding the proper statistical model to estimate the relevant parameters. For a survey of count data estimators, one can see Cameron and Trivedi (1986) and also Gouriéroux and Montfort (1989c). These authors suggest that the extensive statistics literature on stochastic discrete phenomena which proved to be so helpful in biometrics where the theory itself often suggests using a particular statistical model, cannot be applied too directly in econometrics (see, in particular, Cameron and Trivedi). There are good reasons for such a situation. In econometrics and "sociometrics" in general, the data generating process involves rational agents interacting, competitively or cooperatively, over time under constraints of different types such as technological constraints, economic (or market, supplies and demands) constraints, social constraints (tradition, social norms), physical constraints (anatomy, time, space), and so on. These competitive and cooperative interactions give rise to a social or economic equilibrium, a particular set of allocations to and actions by each agent, at which the system is in some sense at rest, typically a dynamic rest (stationarity, ergodicity). Hence, interdependence and simultaneity are very important characteristics of economic outcomes and therefore important concepts in economic theory and applied econometrics; in other words, the future is never really determined until it has fallen into the past! In such a context, economic data generating processes will rarely conform to standard statistical models and distributions: at best, it will be "known" that the true distribution – which will rarely exist in closed form – belongs to a certain family. Specification problems must continuously be addressed and robust estimators of parametric and non–parametric models must be developed.

The case of car accidents is, in many respects, typical of such problems [Greenwood and Yule (1920), Seal (1969), Johnson and Hey (1971), Ferreira (1974) and Lemaire (1985)]. Most accidents involve directly more than one individual, usually car drivers, but also pedestrians and cyclists. Each of those individuals must be assumed or even postulated to behave rationally (the economic notion of rationality rests basically on the transitivity of "constrained" choices and not on their morality or social acceptability; moreover, it may be rational to express "irrationality" or "bounded rationality") and make coherent choices over available possibilities in (any period of) time and over time. Car accidents are the results, i.e. elements of a resulting equilibrium, of interactions between the individual drivers, pedestrians and cyclists. Clearly, such a data generating process is far from the natural, experimental or biological inanimate processes which lead to most statistical models around. Economic and social theories of data generating processes are structural representations of individuals' information, constraints and behavior interacting with each other in a very complex way. These structural representations may or may not have simple reduced form counterparts whose parameters (identifiable, unidentifiable or partly identifiable mixtures of structural parameters and constants) are to be estimated.

This paper deals almost exclusively with reduced form equations relating car accidents to individual characteristics seen as exogenous variables, but we will discuss in section III the use of our econometric results for the determination of an optimal pricing of car insurance, a very delicate calibrating exercise aimed at reducing (or increasing) the rate or number of car accidents to its socially optimal level (or to their socially optimal levels if many different types of accidents, e.g., fatal or not) are considered. Before reaching those complex issues, we will present, in the first section, the data base which will be used here. The particular features of the data base will motivate many ulterior developments and choices of topics.

Subsection II.1 will discuss the econometrics of count data models and more precisely the basic and traditional, although very restrictive, unique (requiring that the mean and the variance be equal) and constant (assuming homogeneous underlying individual car drivers) parameter Poisson distribution; and the similarly important negative binomial distribution, a specific and most interesting two–parameter form of the general Compound Poisson distribution (with a variance different from the mean by allowing the underlying individuals to differ according to their probability of accident, that is to face different one–parameter Poisson distribution). These two distributions are benchmark models of count data and as such interesting ones, although they are unlikely candidates even as reduced form representations of a data generating process as complex as that of car accidents.

Subsection II.2 will discuss related models mentioned above in which the dependent variable, the variable chosen by the individual agents, is continuous (in our case the probability of accident) and therefore not a count variable, but cannot be observed by an outside observer. This observer can only record a proxy variable which is a non-negative categorical variable and therefore an apparent count variable. Four models will be discussed and estimated: the linear probability model, the probit and logit models and the multinomial logit model. In each case, individual characteristics are introduced in the regression. Subsection II.3 will let the moments of the Poisson and negative binomial distributions depend on individual characteristics; therefore, we will assume, as in subsection II.2, that the underlying individuals are heterogeneous in driving skills and attitudes. We will make use of our data bank to define a space of characteristics in which each individual (n = 19 013 of them) will be located. Subsection II.4 will discuss advanced sampling and estimation issues such as endogenous sampling, simulation based inference and pseudo maximum likelihood estimators, three topics of great importance for count data econometrics. In subsection II.5, we will compare the different estimated models by referring to estimated probability of accident of different fictitious individuals. Finally, in section III, we will discuss the potential applications of our results to the determination of an optimal pricing formula for car insurance; of course, our econometric estimates of the link between individual characteristics and the probability of accidents is only one element in such a determination. The conclusion summarizes the main results.

## I. THE DATA

In Québec, the car insurance regime is a mixed system since 1978. Bodily injuries are insured, up to a maximum amount, by a provincial state-owned insurance company, the Régie de l'Assurance Automobile du Québec (RAAQ). Property damages are insured by a competitive system of private insurance companies. The public insurance scheme for bodily injuries is compulsory for all drivers. It is a centrally administered and no-fault system which covers every driver and pays indemnities for medical and rehabilitation costs (without a maximum) and for injuries, dismemberment and pain and loss of enjoyment of life (with a maximum); there is also a maximum indemnity paid in the form of indexed annuities for disability (which represents full coverage for 85% of the population) and death (a pension-type indemnity when there are dependants; a small lump-sum payment otherwise). Although a driver has the option of purchasing additional insurance coverage from private companies (on which we have no data), the no-fault system suspends the right to sue for bodily injuries. Car

owners must also buy a minimum insurance from a private company covering responsibility for property damages to another party (on which we have no data). Everything else is optional.

There are two main types of pricing corresponding to the two main insurance plans: the public plan for bodily injury and the private plan for property damage and collision insurance. The pricing procedure in the former case is very simple. The main sources of financing are from drivers' permits and automobile registration fees (a tax on gasoline was used before March 1982; interest revenues are another source of funds). There is no classification of risks, and drivers pay the same premium whatever their personal characteristics (age, sex, marital status, territory), the car they use, their driving experience and their accident record. Weight and type of vehicle driven are taken into consideration for vehicles other than pleasure vehicles. The private sector uses a traditional classification system for pricing property damage and collision insurance, based on demographic characteristics, automobile use, territory, type of car and past experience for the rating of both property damage insurance and collision insurance. The primary demographic variables are age, sex and marital status. For pleasure vehicles, pricing is based on territory, driver's class, past experience and car group. For other vehicles, driver's classes are not used but the categories of vehicle are important. Driving experience is measured by the number of driving years without accidents with fault since the last accident. Insurance premiums increase with the number of past accidents. For example, if in the previous three years, an individual was responsible for three accidents, the premium is increased by 30% and, for each additional accident, by 10%. Convictions for driving infractions are also used. For example, four convictions for infractions concerning excess speed and other types of traffic violations increase the premium by 25% and for each additional conviction, by 15%. The application of these rules is not always uniform across insurers, however [see Boyer and Dionne (1987b) for more details].

The sample upon which the study is based includes a similar set of data on 19,013 holders of a driving permit on August 1, 1983. The systematic sampling method was applied: first a driving permit was randomly selected on May 16, 1983 among the first 185 in the bank and afterwards each 185th was chosen (once eliminated, all the permits which were temporary, fictitious, expired before August 1, 1980 or belonging to nonresidents). This sample has many advantages. Firstly, it was randomly gathered from among the entire population of Québec drivers and therefore, is likely to be more representative than a sample of insured drivers from a given insurance company. Secondly, it contains five types of information on each driver: the

personal characteristics appearing on the date of the sampling on the driver's license itself (age, sex, experience, place of residence, driving restrictions, class of license) and for the three year period from August 1, 1980 to July 31, 1983 the yearly number of involvements in accidents, the yearly number of demerit points, the yearly number of license suspensions for criminal offenses, and the number of days of validity of the permit during each yearly period. Thirdly, the data is historical and can be updated at minimal costs so as to eventually cover as many years as desired. This is clearly a very important advantage of the data since the probability of accident in any given year is relatively low which may make the possibility of panel data even more valuable. Finally, this data makes it technically feasible to send a questionnaire to a subsample of drivers in order to collect additional information on driving habits, risk exposure, socioeconomic characteristics, and so on. It must be noted, however, that for reasons of confidentiality, the name and address of the individual drivers are not known to us, although a correspondence table is kept at the RAAQ for the purpose of updating the files.

It must be noted also that the data bank also has important and serious limits. The most important is certainly the absence of data measuring directly the individual exposure to risk and for the purpose of using the econometric estimates for determining a fair and efficient pricing formula for automobile insurance, the absence of data on private insurance premia (for property damage and for injuries in excess of the public maximum coverage).

Our data is better described with the help of self-explanatory Tables I.1 to I.6. In those tables, each cell contains three numbers: the first one indicates the number of drivers in the cell, the second one gives that number expressed as a percentage of the row total and the third as a percentage of the column total.

## II. THE ECONOMETRICS OF CATEGORICAL AND COUNT DATA MODELS

### II.1 Count data models without individual characteristics

It is reasonable to suppose that for a given individual, the probability of being involved in a car accident satisfies the following conditions: the probability of accident is proportional to the length of the period considered, the instantaneous probability of accident is constant throughout the period, the probability of having more than one accident in a period is small and accidents are independent. In others words, letting $k(t, t + dt)$ stand for the number of accidents occurring in the time interval from $t$ to $t + dt$, it is reasonable to suppose that

$$p[k(t, t + dt) = 0] = 1 - \lambda dt + O(dt)$$

and that

$$p[k(t, t + dt) = 1] = \lambda dt + O(dt);$$

therefore,

$$p[k(t, t + dt) \geq 2] = O(dt) \text{ when } dt \to 0.$$

The Poisson distribution does satisfy the above characteristics and has therefore been used as a first choice to represent the distribution of accidents. The probability that an individual will be involved in k accidents over t years is then:

$$p(k \mid t, \lambda) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \tag{4}$$

where $\lambda$ is both the mean and variance of the distribution. However, using a Poisson distribution to represent the observed distribution of accidents for a group of individuals supposes implicitly that all individuals have the same probability of being involved in one or more accidents. This last feature is very restrictive. If, on the other hand, we assume that $\lambda$ varies between individuals and that for a given individual the distribution of accidents follows a Poisson distribution, then it will be mathematically convenient to suppose that $\lambda$ is distributed as $f(\lambda)$,

$$f(\lambda) = \frac{\lambda^{a-1}\tau^a e^{-\tau}}{\Gamma(a)}, \quad \tau > 0 \text{ and } a > 0 \tag{5}$$

that is $\lambda$ follows a Gamma density function with mean $a/\tau$ and variance $a/\tau^2$, $\Gamma(a)$ being the Gamma function. From the distribution of $\lambda$, the *average* probability of being involved in k accidents over t years or periods becomes the negative binomial distribution:

$$\begin{aligned} P(k \mid t, a, m) &= \int_0^\infty P(k \mid t, \lambda) f(\lambda) d\lambda \\ &= \frac{\Gamma(k+a)}{k!\Gamma(a)} \left[\frac{a}{a+mt}\right]^a \left[\frac{m\,t}{a+mt}\right]^k \end{aligned} \tag{6}$$

where the mean is $m = a/\tau$ and the variance is $(a/\tau)(1 + 1/\tau)$ [see Ferreira (1974) and Lemaire (1985) for more details]. If the distribution of accidents is better represented by the Poisson distribution (4), it would suggest that individuals are homogeneous insofar as their respective

probabilities of accident are concerned. This would imply among other things that differential insurance pricing would have no supporting empirical evidence. We estimated the two distributions (4) and (6) for the 12–month period from August 1982 to July 1983 by the maximum likelihood method. Results are reported in Table II.1, where k is the number of accidents, $n_k$ is the number of individuals with k accidents and LL is the log–likelihood. For the Poisson distribution, we obtain an estimated $\hat{\lambda}$ of 0.0701; observed and predicted values for $n_k$ give a $\chi^2$ statistic of 133.06 thereby rejecting the hypothesis that the distribution of accidents follows a Poisson distribution. As for the negative binomial distribution, the maximum likelihood estimates are $\hat{a} = 0.6960$ and $\hat{\tau} = 9.9359$ and therefore $\hat{m} = 0.0701$; observed and predicted values for $n_k$ give a $\chi^2$ statistic of 2.21. We cannot therefore reject the hypothesis that drivers are indeed heterogeneous insofar as their respective probabilities of accident are concerned. These results provide the necessary empirical justification for proceeding with the econometric estimation of the probability of accident by controlling for the individual characteristics in order to isolate properly the specific effects of different individual characteristics. It is interesting to note that based on the parameter estimates of the negative binomial distribution, equation (6) gives the probability of having 0 accident (k = 0) in 1982–83 (t = 1) as 0.935 and therefore

$$P(k \geq 1 | t = 1, a, m) = 0.065.$$

This value corresponds to the mean of the estimated individual probabilities of accident to be obtained below from the probit and logit models.

### II.2 Categorical models: linear probability, probit, logit and multinomial logit models

The standard econometric models used to estimate probabilities of accident from the characteristics of individuals and with variables indirectly representing the individual's choice of safety activities are categorical models where the dependent variable, the number of accidents, is a binary variable (0 for no accident and 1 for a positive number of accidents) or a multinary (rarely more than three categories) variable. We will consider here the main models: the standard linear probability model estimated with the ordinary least squares estimator and the nonlinear probit, logit and multinomial logit models; the first three are binary models and the fourth is a trinary model. In each case, a first group of variables includes age, sex, number of years since the first license was obtained, place of residence, driving restrictions, class of the driver's license, and period of validity of the license; we will represent those by the vector $v_i$. Each variable, except for the last one, is represented by a set

of dichotomous variables, as is customary the case in the insurance business. For example, the "age" variable is represented by eight dichotomous variables with the age group 16–19 in the constant. The exact meaning and measure of each variable are presented in Table A.1 (Appendix). A second group consists of proxy variables for the unobservable choices of self–protection activities by the individual; we will represent those by the vector $w_i$. It corresponds to the individual's past driving experience. Three such variables appear in our sample: the cumulative number of demerit points D over the two previous years, the number of license suspensions S for criminal offenses in the previous year (we have no data on this for 80–81) and, finally, the number of accidents Z the individual was involved in during the two previous years.

The first model is the linear probability model:

$$k_i = \alpha + \beta X_i + u_i \tag{7}$$

where $X_i = (v_i, w_i)$ and $E(u_i) = 0$. The dependent variable $k_i$ is dichotomous: $k_i = 1$ if the i–th individual had one or more accidents during 1982–83, and $k_i = 0$ otherwise. $E(k_i | X_i) = \alpha + \beta X_i$ can be interpreted as the probability that $k_i = 1$ since

$$E(k_i | X_i) = p(1 | X_i)1 + (1 - p(1 | X_i))0 = p(1 | X_i). \tag{8}$$

Table II.2 presents the results of the application of ordinary least squares to this model. The first column (Model 1) corresponds to only one proxy for care activities, namely the Z variable; the second column (Model 2) corresponds to all three proxies D, S, Z. Values in parentheses are t–statistics.

It is interesting to note that the age variables are significant. In comparison with the reference group (i.e. the 16–19 year–old drivers), we see that drivers over 25 will have a lower probability of accident by 3.3 to 5.8 percentage points in the first model, and by 3.0 to 4.7 percentage points in the second model. Sex is also significant, male drivers having a probability of accident 4.4 and 3.4 points higher respectively.

As was mentioned above, these variables can also be proxies for driver skill and, more importantly, risk exposure. Our results are similar to those obtained by Dahlby (1983) and somewhat justify their use by insurers. Clearly, individual measures of risk exposure and driver skill would be more appropriate here. However, variables VALA, VALB and VALC measure the number of days during which the permit is valid. Those variables indirectly take

into account the fact that a new driver may not have a full year of driving experience and also the fact that individuals with a license suspension do not drive during that time. We observe that the effect of VALC is positive and significant. The place of residence variable is not significant except for one region, meaning that when individual information is available, this variable, used as proxy of risk by insurers, is irrelevant. Finally, driving licenses for classes 11 and 12 (bus drivers), 21 and 22 (truck and trailer truck drivers), 31 (taxi drivers) and 55 (small motorcycles) are significant, with classes 11 and 12, 21 and 22, and 31 raising the individual probabilities of accident and class 55 reducing them.

Our results also indicate that the above variables explain only in part the differences in probabilities between drivers. Indeed, the three "behavioral" variables D "demerit points", S "license suspensions for criminal offense" and Z "past accident involvements" are significant [the correlation coefficients between D, S and Z are $r_{DS} = 0.07901$, $r_{DZ} = 0.22110$ and $r_{SZ} = 0.09431$]. In the first model, a single accident during 1980–81 or 1981–82 raises the probability of being involved in an accident in 1982–83 by 4.8 percentage points. Considering that the average probability of being involved in at least one accident is 6.5%, the increase is quite important. In the second model, each additional demerit point increases the probability of accident by 1.1 percentage point; an additional license suspension increases the probability of accident by 5.7 percentage points; and an additional involvement in an accident during 1980–81 or 1981–82 increases the probability of accident in 1982–83 by 4.0 percentage points. Although the regressions are generally very significant ($F = 13.8$ and 16.5), the adjusted $R^2$ ($\bar{R}^2$) are only 0.03 and 0.04. This result was expected since the data is at the individual level and therefore very disaggregated, and since the event "accident" is rather rare [Theil (1971), p. 181]. [A regression on the data aggregated over 353 postal codes gave an adjusted $R^2$ of 0.76, with the three variables D, S, Z also being quite significant.] We can choose between the two models by testing the $\beta_D = \beta_S = 0$ hypothesis in model 2. We obtained an F-statistic of 82.06 well above the $F^{99}_{(2,\infty)} = 4.60$ bound and model 2 therefore appears significantly more adequate.

We made a thorough investigation of multicollinearity. We ran auxiliary regressions of X, Y and Z on all the other explanatory variables. We obtained low $R^2$'s (0.13, 0.09 and 0.05 respectively). However, those regressions are not appropriate if there exist other relations between the explanatory variables. We therefore used the approach proposed by Belsley, Kuh and Welsch (1980) [referred to as BKW]. This method based on the analysis of the

characteristic roots and vectors of the matrix X'X (scaled to unit length but not centered) can reveal the presence and perhaps the nature of multicollinearity. We found three near linear dependencies associated with "condition indexes" greater than 30 (the heuristic rule proposed by BKW says that condition indexes of 30 to 100 are associated with moderate to strong relations). The examination of the variance–decomposition proportions indicated competing dependencies involving the constant, the age variables, the number–of–years–with–a–driving–license variables, and the validity variables. Auxiliary regressions were performed as suggested by BKW (pp. 150–160) for the validity variables on all the other explanatory variables; we obtained $R^2$s of 0.87, 0.55 and 0.07 respectively for VALA, VALB and VALC, which indicate that the relations are not very tight. These results indicate that the multicollinearity problem is not harmful in our case.

However, the linear regression model is far from being totally adequate when the dependent variable is dichotomous. Firstly, there is a serious heteroscedasticity problem and secondly, $\alpha + \beta X_i$ could possibly be outside the range [0,1] for certain $X_i$; this is troublesome if we interpret $E(k|X_i)$ as a probability. Although these problems are likely to be less serious given the large number of observations and variables, we obtained $\alpha + \beta X_i < 0$ for 602 individuals (3.2 % of our sample). More adequate nonlinear models such as the probit and the logit models should be estimated.

Let us consider again the linear model $y_i = \alpha + \beta X_i$ with the dependent variable being the probability of accident. This variable is not observed as such. What is observed is not the probability of accident but rather whether the individual had an accident or not. We may pose that the occurrence of an accident depends on the probability of accident itself and random elements which can be represented as $u_i$: hence, an individual has an accident if $p(\text{accidents}) + u_i = \alpha + \beta x_i + u_i > 0$ and no accident if $u_i < -p(\text{accident})$. Therefore, even if $\alpha + \beta X_i$ does represent the probability of accident for driver i, we observe only that this driver had at least one accident ($k_i^* = 1$, i.e. $u_i > -\alpha - \beta X_i$) or had no accident ($k_i^* = 0$, i.e. $u_i < -\alpha - \beta X_i$). Therefore, the likelihood function will be given by:

$$LF = \left[\prod_{k_i^*=0} F(-\alpha - \beta X_i)\right]\left[\prod_{k_i^*=1} (1 - F(-\alpha - \beta X_i))\right] \qquad (9)$$

where $F(\,)$ is the cumulative distribution function of $u_i$. If $u_i$ is normally distributed, an assumption which could be justified by the fact that the number of factors which influence the probability of accident other than those factors included in $X_i$ is indeed very large, then

$$F(-\alpha - \beta X_i) = \int_{-\infty}^{\frac{-\alpha-\beta'X_i}{\sigma}} (2\pi)^{-1/2} \exp(-\frac{s^2}{2}) ds \qquad (10)$$

and the maximization of LF with respect to $\alpha$ and $\beta$ (setting $\sigma = 1$, because LF is homogeneous of degree 0 in $\alpha$, $\beta$ and $\sigma$) gives the probit estimator. Probit estimates $\alpha$ and $\beta$ for Model 1 and Model 2 are presented in Table II.3 under the columns "Original Coefficient". As a whole, the regressions are very significant. Moreover, the sets of significant variables are basically the same as those obtained with OLS. Since the probit estimator is nonlinear in the parameters, we observe that the average individual, defined by the vector of average values of the explanatory variables, has an estimated probability of accident of 0.053 in Model 1 and 0.052 in Model 2. These are significantly smaller than the mean of the estimated probabilities of accident, which is 0.065 in both models, over the whole sample of 19 013 observations. This mean estimated probability differs also from the mean 0.071 of the previously estimated negative binomial distribution. The difference is due to the use of a cardinal count variable having values 0, 1, 2, 3, etc. to measure the number of accidents in the negative binomial model while the dependent variable in the probit model is an ordinal categorical 0 – 1 variable (k = 1 for one or more accidents).

We performed a log-likelihood ratio test of the $\beta_D = \beta_S = 0$ hypothesis in model 2 to discriminate between the two models; we obtained $-2(LL_1 - LL_2) = 95.40$ well above the $\chi^2_{2,99} = 9.21$ bound, again favoring significantly model 2 over model 1. This result corroborates the result of a similar test in Table II.2 and, from now on, we will therefore concentrate on model 2 with variables D, S and Z.

Table II.3 also presents transformed coefficient estimates which can be directly interpreted as the variation in the probability of accident due to the presence of a given variable. Consider a variable $w_k$ and its sample mean $W_k$. The transformed estimate of the coefficient for $w_k$ is the difference between the probability of accident (10) estimated for ($w_k = 0$, $w_j = W_j$ for $j \neq k$) and its value estimated for ($w_k = v_k$, $w_j = W_j$ for $j \neq k$), where $v_k$ equals 1 for a dichotomous variable and may take different step values for other variables. As an example of the former, consider the results of model 2: the otherwise average male driver (SEXM = 1) has a probability of accident 3.7 percentage points higher than the otherwise average female driver (SEXM = 0). As examples of the latter, consider again the results of model 2: the otherwise average driver with five demerit points in the previous two-year

period (D = 5) has a probability of accident which is 3.4 (0.6+2.8) percentage points higher than that of the otherwise average driver with no demerit points and, similarly, the otherwise average driver with two previous accidents (Z = 2) has a probability of accident which is 3.4 points larger than that of the otherwise average driver with only one previous accident (Z = 1) and 5.9 (2.5 + 3.4) points larger than that of the otherwise average driver with no previous accidents (Z = 0).

It is interesting to note also that age is quite significant, drivers over 25 years of age having probabilities of accidents 2.0 to 3.3 percentage points lower than that of the 16-19 year-old drivers. Experience, or more precisely the number of years with a driving license, is not very significant; the same is true of the region of residence with one exception. It appears that restriction A (must wear glasses or correcting lenses) reduces significantly the probability of accidents, possibly because some (many) drivers who do not have such a restriction should in fact have it! On the other hand, drivers who can only drive automatic transmission equipped cars and drivers who have a temporary 6-month license are risky drivers. Finally, truck and trailer truck drivers and taxi drivers are significantly more risky drivers (they may also have more risk exposure).

Tables II.4 to II.7 provide estimated probabilities of accident of the otherwise average individual as a function of cumulated demerit points, suspensions and past involvements in accidents. As one can see, those "behavioral" variables raise quite sharply the probability of accident.

We saw above that male drivers are more risky and that young drivers (less than 25) are also more risky. What about young female drivers? We looked at interaction effects between age and sex, a debated issue in the car insurance context. We obtained probit estimates for the age*sex interaction terms as follows. We constructed categorical variables F19 (females 19 years of age or younger on August 1, 1980; the omitted category), M19 (males 19 years of age or younger), F2024, M2024, F2534, M2534, etc. The results are:

**Probit Estimates with Age*Sex Interaction Variables**

| VARIABLE | PROBIT ESTIMATE (t–statistics) |
|---|---|
| M19 | 0.478 (4.24)** |
| M2024 | 0.443 (3.65)** |
| M2534 | 0.218 (1.75)* |
| M3544 | 0.166 (1.28) |
| M4554 | 0.165 (1.23) |
| M5564 | 0.086 (0.62) |
| M65 | 0.021 (0.14) |
| F2024 | 0.033 (0.26) |
| F2534 | –0.058 (–0.46) |
| F3544 | –0.224 (–1.63) |
| F4554 | –0.209 (–1.42) |
| F5564 | –0.569 (–2.91)** |
| F65 | –0.004 (–0.02) |
| D | 0.056 (9.62)** |
| S | 0.291 (2.23)** |
| Z | 0.211 (8.11)** |
| CONST. | –2.382 (–10.04)** |

(** significant at 95%; * significant at 90%)

These results suggest the following comments: although the age coefficients for females have the expected negative sign, only the coefficient of F5564 is significant; as for males, the sex factor dominates the age factor as expected and we note that the probability of accident decreases with age but only the coefficients of the younger groups are significant; moreover, the reduction in the probability of accident happens sooner for males than for females but becomes more important at age 55–64 for females than for males; finally and most importantly, the coefficients of variables D, S and Z were unaffected (all the other coefficients not reported here were also unaffected).

The results of Table II.3 show that the number of demerit points D and the number of license suspensions for criminal offences S cumulated over time are very significant to explain the individual probabilities of accidents. These two variables are linear aggregates in the sense that they give the same weight to the number of demerit points corresponding to the different infractions or offences described in Tables A.2 and A.3. In a recent study [Boyer, Dionne and Vanasse (1988)], we investigated two questions related to the above results and form of aggregation. Firstly, if we disaggregate those variables and use different variables for different infractions and offences, which infractions and offences in Tables A.2 and A.3 would have a larger effect on the individual probabilities of accidents. Secondly, are the relative

number of points for each infraction presented in Table A.2 and obtained from the actual regulation of road safety in Québec adequate? In other words, do they represent the marginal effect of those infractions on the individual probabilities of accidents? Our results, available upon request, indicate that the offences for driving under the influence of alcohol are more significant than the other offences of Table A.3 and that among the driving infractions of Table A.2, the not stopping at a stop sign and not stopping at a red light infractions are not enough penalized (in the sense of carrying too few points) relatively to the excess speed infractions even if both are significant in explaining the individual probabilities of accidents. Moreover, our results show that the effect of each infraction is nonlinear which means that the adequate number of points for each infraction should be an increasing function of the number of infractions.

As for the linear probability model, it is desirable to check for the presence of multicollinearity in the probit model. Even if the methodology used above for the analysis of multicollinearity in the linear model is not necessarily appropriate for nonlinear models such as the probit, we nevertheless undertook some analysis. The extension of the methodology to the probit model is far from being straightforward. The estimates of the probit model are the solution of

$$\frac{\partial LLF}{\partial \beta} = \sum_{i=1}^{n} \frac{K_i - F(X_i\beta)}{F(X_i\beta)[1 - F(X_i\beta)]} f(X_i\beta)X_i' = 0, \tag{11}$$

where LLF is the log–likelihood function. The estimated variance–covariance matrix of the parameters is

$$[E(-\partial^2 L/\partial\beta\partial\beta')]^{-1} = (X'\Lambda X)^{-1}. \tag{12}$$

where $\Lambda$ is a diagonal matrix whose element (i,i) is

$$\frac{f^2(X_i\beta)}{F(X_i\beta)[1 - F(X_i\beta)]}.$$

As pointed out by BKW, multicollinearity can lead to convergence problems (we have none) if the optimization algorithm uses the inverse of the Hessian matrix and can lead also to poor estimates of the variances of the parameters. We applied the same methodology to the matrix $X'\Lambda X$ (instead of $X'X$) and found the same kind of relations involving the same groups of variables as in the linear case. We then re–estimated the probit model without the

number–of–years–with–a–driving–license variables and the validity variables. We also applied the BKW method to the matrix of explanatory variables and the Hessian matrix of the "reduced" model and we found no evidence of multicollinearity. We obtained results very similar to those of the complete model except for the constant and the age variables. These results lead us to the conclusion that multicollinearity is not harmful in our case and that low multicollinearity is preferable to a misspecification of the model due to the omission of important variables such as the number–of–years–with–a–driving–license variables. For example, when calculating insurance premiums, it seems important to make the difference between the 20–year old driver with a 2–year old license and the 24–year old with a 6–year old license. Testing the "reduced" specification against the "full" specification by setting the null hypothesis ($H_0$: the coefficients of the number–of–years–with–a–driving–license variables and of the validity variables all equal to 0) lead us to reject the null hypothesis (likelihood ratio test statistics equal to 63.4). Therefore, there is a gain in the ability of the model to predict the probability of accident when the "full" specification is used. Another fact is that the predictions will not necessarily be imprecise because of multicollinearity as long as the values of the independent variables for which a prediction is desired obey the same near exact restrictions as the X matrix. For all these reasons, we think that it is legitimate to use the "full" specification model.

The probit model was obtained above by supposing that $F(-\alpha - \beta X_i)$ in (9) is the cumulative normal distribution. One may wish to verify if the estimated coefficients of the explanatory variables are sensitive to that assumption. A standard alternative is to suppose that the random elements $\mu_i$ follow a Wiebull distribution which has the interesting property that its cumulative distribution can be written directly as

$$F(-\alpha - \beta X_i) = \frac{\exp(-\alpha - \beta X_i)}{1 + \exp(-\alpha - \beta X_i)} = \frac{1}{1 + \exp(\alpha + \beta X_i)} \qquad (13)$$

which substituted in (9) leads to the logit estimator. The logit estimates are given in Table II.4. To compare the logit estimates with the probit estimates of Table II.3, we must first divide the logit estimates by $(\pi/\sqrt{3})$ to bring their variance to a comparable level with the probit estimates. Recall that the probit likelihood function is from (9) and (10) homogeneous of degree 1 in $\alpha$, $\beta$ and $\sigma$ which lead us to pose $\sigma = 1$; but in the case of the logit estimator, $\sigma = \pi^2/3$. The adjusted logit coefficient estimates are given in the last column of Table II.4. First, the significant variables are the same and second, the adjusted logit coefficients are

quite close to their probit counterpart especially so for the more significant variables; the only noticeable difference is for the age and sex variables for which the logit estimates are about 20 % larger (in absolute value). Indeed, with 19 013 observations, one could have expected even larger differences [Maddala (1977), pp. 23-24]. We can therefore be reasonably confident in either the probit or the logit estimates but since the probit model appears slightly better, we will use only the probit estimates in the comparison of models in section II.5 below.

An important drawback of both the probit and logit models is that the dependent variable is restricted to 0 (if no accident) and 1 (if one or more accidents). Clearly, it may be interesting to consider more than two categories in the case of car accidents in order to better represent the tail of the distribution. Given that only 11 individuals in our sample had more than two accidents in 1982-83, we will consider a multinomial model with three categories: no accidents, one accident, two and more accidents containing respectively 17 784, 1 139 and 90 individuals. Considering the relative computational complexity and cost of multinomial probit and logit models, we proceeded to the estimation of the multinomial logit model only. However, we should emphasize that the consideration of one more category (2+) implied the estimation of 53 additional parameters.

The transformed coefficients appearing in Table II.5 give, as in Table II.3, the impact of a variable on the probability of accident for the otherwise average individual. They were obtained as in Table II.3 except that we now have two probabilities: the probability of having exactly one accident Prob1 and the probability of having two and more accidents Prob2+. Consider, for example, the SEXM variable: the otherwise average male driver (SEXM = 1, $w_k = W_k' \; \forall \, k \neq SEX$) has a probability of having one accident which is 3.3 percentage points higher than the otherwise average female driver (SEXM = 0, $w_k = W_k \; \forall \, k \neq SEX$) and a probability of having 2+ accidents which is 0.2 percentage points higher; considering that the average probabilities are respectively 4.74 % and 0.19 %, the relative variations are quite large in this case. Similarly, the otherwise average individual who had one license suspension for criminal offense in 1981-82 (S = 2) has a probability of having one accident in 1982-83 which is 4.78 percentage points higher that the otherwise average driver who had "only" one such suspension (S = 1) and a probability of having two accidents or more in 1982-83 which is 0.33 percentage points higher; again, these relative increases are quite large.

The above probit and multinomial logit models are categorical models and therefore may indeed miss the basic cardinal aspect of car accidents. In a sense, having two accidents is not simply being in a category of those with two or more accidents and not in the category of those with one accident, but it means that one had twice as many accidents as those in the

one-accident category. It is also quite different from purely ordinal categorization, such as riding a bus or a train, given that one does not use a private car. In other words, car accidents represent a count process where numbers 0, 1, 2, 3, etc. are not simply categories but truly cardinal numbers which can be added and compared. Moreover, count data models will naturally discriminate between those with, say, 3 and 4 accidents by using all the information contained in the data. Hence, it seems desirable to look at count data models with individual characteristics which we will do in the next subsection.

### II.3 Count data models with individual characteristics in the regression component

The constant individual probability of accident models of subsection II.1 are clearly quite restrictive. The favorable $\chi^2$ test of the negative binomial model in Table II.1 suggests that individual car drivers are likely to have different probabilities of accident. Hence, it would be worthwhile to find out which individual characteristics are responsible for such heterogeneity. Moreover, allowing heterogeneity in individual probabilities of accident may make the basic Poisson model more attractive. In this subsection, we look at both the Poisson and the negative binomial models with heterogeneous car drivers. In both cases, we will assume that an individual's probability of accident depends on his personal characteristics: in the Poisson model, we assume that the distribution of accident an individual is involved in remains a one-parameter distribution (hence, mean and variance are equal), while in the negative binomial case, we allow for over-dispersion (a variance larger than the mean) in the distribution by assuming that the individual characteristics leave some differences between the individual probabilities unexplained and that this unexplained portion follows a Gamma distribution. In a sense, the set of characteristics which the data allow us to define may not be rich enough to stick to the Poisson distribution with heterogeneous drivers.

Let $K_i$ be the number of accidents in which an individual i is involved in a given period. Let us suppose that the assumptions underlying the one-parameter Poisson model are satisfied and moreover, that the function linking the parameter $\lambda_i$ to the individual's characteristics is

$$\lambda_i = \exp(X_i\beta)$$

where $\beta$ is a k x 1 vector of weights to be estimated. Then, the probability that the individual i will be involved in k accidents is

$$p(K_i=k \mid X_i) = \frac{e^{-\exp(X_i\beta)}\exp(X_i\beta)^k}{k!}.$$

The mean and variance of $K_i$ are both $\lambda_i = \exp(X_i\beta)$; note that the form $\exp(\cdot)$ simply makes sure that $\lambda_i$ is non-negative. The likelihood function of the data $(K_i, X_i)$ is given by

$$L(K; \beta) = \prod_{i=1}^{n} \frac{e^{-\exp(X_i\beta)}\exp(X_i\beta)^{K_i}}{K_i!}$$

The estimates of $\beta$ are given in Table II.6, together with transformed coefficents which can be interpreted as in Table II.5.

The above Poisson model with individual characteristics is restrictive in many ways. As we mentioned in subsection II.1, the Poisson model rests on an assumption of independence of successive accidents. Hence, the probability that an individual will be involved in a car accident in period t is independent of whether or not he was involved in an accident in t - 1 or more generally $t' < t$. There are good reasons to believe that in the context of car accidents, this will not be the case. An accident may make someone more nervous on the road at least for some subsequent periods and increase or decrease the probability of accident; an accident may have occurred because the individual is in a state of intensive stress which will last more than one basic period. Those are only two of the reasons one can think of why car accidents may occur in spells, that is possibly linked to some form of dynamic dependency. In fact, our results indicate that past involvements in accidents raise the probability of accidents significantly in a given period. This result may be due to missing individual characteristics which would be significant in explaining the probability of accidents for an individual; but they may indicate the existence of spells of accidents, that is a more fundamental breakdown in one of the basic assumptions of the model, the assumption that events are independent over time. We did not pursue the problem here [see, however, Heckman and Borjas (1980)].

Another possibly restrictive assumption of the Poisson model is that for a given individual, the mean and variance of the distribution be equal. The problem here is of a different nature from the previous one: it is basically an estimation problem since it may produce an underestimate of the variance (towards the mean) and hence too large student-t values, that is too many and too strongly significant individual characteristics. Moreover, the Poisson model may attach two small probabilities to the possibility that the number of

accidents is two or more. There are fortunately relatively simple remedies for this second problem. The idea is, of course, to break the strong link between mean and variance. We noted in the previous subsection that one could generate a two-parameter distribution of accidents by allowing the individual $\lambda_i$ to follow a gamma distribution. And it turned out that allowing for a larger variance improved the results very significantly. In the present case, although the $\lambda_i$ are different across individuals, it remains that the mean and variance of each individual Poisson distribution are both equal to $\lambda_i = \exp(X_i\beta)$. Of course, allowing the $\lambda_i$ to differ across individuals through the different vector $X_i$ of characteristics may or may not be sufficient to explain adequately the overdispersion uncovered in the preceding subsection.

One way to introduce additional flexibility in the model is to suppose that the vector of characteristics X is not sufficient to capture the whole difference between individuals and to assume that the additional unobserved characteristics can be modeled as a random addition $\mu_i$ to $X_i\beta$, that is

$$\lambda_i = \exp(X_i\beta + \mu_i)$$

with probability density function $h(\mu_i)$. Of course, $\mu_i$ may also represent pure randomness and not missing characteristics, that is not a specification error. Then, the marginal probability that the i-th individual will be involved in k accidents becomes:

$$\int P[k \mid t = 1, X_i, \mu_i] h(\mu_i) d\mu_i$$

$$= \int \frac{e^{-(X_i\beta+\mu_i)} \exp(X_i\beta + \mu_i)^k}{k!} h(\mu_i) d\mu_i$$

which is the general form of the Compound Poisson distribution: different specification of $h(\mu_i)$ leads to different particular cases.

The particular form we will consider here is the negative binomial distribution obtained by writing $\lambda_i = \exp(X_i\beta + \mu_i) = \exp(X_i\beta)\varepsilon_i$ and assuming that $\varepsilon_i = \exp(\mu_i)$ follows a gamma density

$$f(\varepsilon) = \frac{\varepsilon^{a-1} e^{-a\varepsilon} a^a}{\Gamma(a)}$$

with a mean of 1 (the mean of $\mu_i$ is assumed equal to 0) and a variance $\frac{1}{a}$. Since $\lambda_i = \exp(X\beta_i)\,\varepsilon_i$, it will have a mean of $\exp(X_i\beta)$ and a variance of $\frac{1}{a}\exp(X_i\beta)^2$. Integrating the expression for the distribution of accidents above, we find

$$P[K_i = k \mid X_i] = \int \frac{e^{-\exp(X_i\beta)\varepsilon_i}[\exp(X_i\beta)\varepsilon_i]^k \varepsilon_i^{a-1} e^{-a\varepsilon_i} a^a}{k!\,\Gamma(a)}\, d\varepsilon_i$$

$$= \frac{\Gamma(k+a)}{k!\,\Gamma(a)}\left[\frac{\exp(X_i\beta)}{a}\right]^k\left[1+\frac{\exp(X_i\beta)}{a}\right]^{-(k+a)}$$

which is the expression for the negative binomial distribution with parameters a and $\exp(X_i\beta)$: its mean is $E(K_i) = \exp(X_i\beta)$ and its variance $V(K_i) = \exp(X_i\beta)[1 + \exp(X_i\beta)/a] = E(\lambda_i)[1 + E(\lambda_i)VAR(\varepsilon)]$. Hence, the variance is a strictly convex increasing function of the mean: the variance/mean ratio is a linear function of the mean. As mentioned by Cameron and Trivedi (1986), this is just one parameterization of the negative binomial distribution. We can rewrite $P[K_i = k]$ above more generally and directly as a two-parameter $(\nu_i, \phi_i)$ distribution

$$P[K_i = k] = EP[K_i = k \mid \lambda_i]$$

$$= \frac{\Gamma(k+\nu_i)}{\Gamma(k+1)\Gamma(\nu_i)}\left[\frac{\nu_i}{\nu_i+\phi_i}\right]^{\nu_i}\left[\frac{\phi_i}{\nu_i+\phi_i}\right]^k \quad ; \phi_i > 0, \nu_i > 0$$

with $E[K_i] = \phi_i$ and $VAR(K_i) = \phi_i + \frac{1}{\nu_i}\phi_i^2$, where both $\nu_i$ and $\phi_i$ are functions of the individual characteristics $X_i$. Since we wish to stick to the interpretation $E(\lambda_i) = \exp(X_i\beta)$, we can set $\phi_i = \exp(X_i\beta)$. But we have some freedom in setting the relation between $\nu_i$ and $X_i$. For instance, we could assume that

$$\nu_i = \frac{1}{r}[\exp(X_i\beta)]^q \; ; r > 0.$$

The above formulation for $P[K_i = k \mid X_i]$ corresponds to the special case $r = \frac{1}{a}$ and $q = 0$. Cameron and Trivedi do in fact consider two cases: $q = 0$ and $q = 1$. We consider only the case $q = 0$ here and hence the heterogeneity of the individual car drivers, composed of a fixed part $\exp(X_i\beta)$ and a random part $\mu_i$ or $\varepsilon_i$, allows for overdispersion in the data but a restricted

form of overdispersion, namely that the variance of the distribution of accidents for individual i is a second-degree polynomial of the mean of the distribution: $V = E + gE^2$ where $g = \frac{1}{a}$ is the variance of $\varepsilon_i$, the random (or missing) term of the individual vector of characteristics. As $g \to 0$ (that is $a \to \infty$), we obtain the Poisson distribution as expected since then $VAR(\varepsilon_i) = 0$. The choice of parameterization of the negative binomial distribution is by no means a simple matter: Cameron and Trivedi (1986) do stress that they correspond to different models of the underlying data generating process. We feel that assuming $q = 0$ above is a very reasonable assumption to make in the case of car accidents. Drivers who on average are worse drivers, that is have more accidents, tend to experience not only larger variance but also a variance/mean ratio which is about 10.9% higher for the groups who had 2 or 3+ accidents in 1980-1982 compared to the groups who had 0 or 1 accidents in 1982-1983. This corresponds rather closely to the estimated value $(\hat{\tau})^{-1} = 10.06$ in Table II.1 if we pool together the first two and the last two groups of the table. However, this gives us only two points, a rather limited basis for assessing the shape of the function relating $(\frac{V}{E})$ and E!

| Nb of Accidents in 1980-1982 | Nb of Individuals | Nb of Accidents in 1982-1983 | Mean E | Variance V | V/E | |
|---|---|---|---|---|---|---|
| 0 | 16,349 | 963 | 0.059 | 0.064 | 1.085 | } 1.083 |
| 1 | 2,213 | 253 | 0.114 | 0.122 | 1.070 | |
| 2 | 364 | 81 | 0.223 | 0.263 | 1.179 | } 1.177 |
| 3+(= 3.182) | 87 | 35 | 0.402 | 0.470 | 1.169 | |

Coming back to the estimation of the parameters of $P[K_i = k | X_i]$, we should first note that this model or this particular parameterization of the negative binomial model is equivalent to that of Gouriéroux, Montfort and Trognon (1984b), of Haussman, Hall and Griliches (1984) and to the Megbin II model of Cameron and Trivedi (1986). The maximum likelihood estimates are given in Table II.7 together with transformed coefficients which can be interpreted as in Tables II.5 and II.6.

The Poisson and negative binomial estimates are rather close and the transformed coefficients are similarly quite comparable. Moreover, the same variables are significant. Since the Poisson is a particular case of the negative binomial model - it corresponds to the restriction $(\frac{1}{a}) = 0$ on the negative binomial model -, we can perform a log-likelihood ratio test

of this restriction to decide if considering individual characteristics is sufficient to explain the overdispersion in the data or if we gain in adding a gamma-distributed random element to $X\beta$. We obtain -2(LL(Poisson) - LL(negative binomial)) = -2(-4563.86 + 4556.43) = 14.86 which is larger than $\chi^2_{1,99} = 6.63$ indicating that the negative binomial model does perform significantly better. Finally, the log-likelihood ratio test indicates that the standard negative binomial model estimated without individual characteristics is rejected when compared with the negative binomial model estimated with individual characteristics [-2(-4,916.78 + 4,556.43) = 720.7 while the critical value at a 5% significance level is 63.83].

### II.4 Related sampling and estimation issues

Maximum likelihood estimators (MLE) are not always efficient or easily tractable and their properties are quite sensitive to the distributional assumptions and to the sampling process. Most importantly, MLE of the Poisson or negative binomial regression models will not be efficient if the "true" distribution is not the specified one (this also holds for other regression models). Gouriéroux, Monfort and Trognon (1984a, GMT hereafter) show that, if the "true" distribution is a member of the linear exponential family (LEF), pseudo maximum likelihood methods will be convergent, provided that the mean of the distribution is well-specified. Since Poisson and negative binomial distributions are members of the LEF, pseudo maximum likelihood methods can be used to estimate these regression models. The pseudo maximum likelihood functions will be

$$-\sum_{i=1}^{n} \exp(X_i\beta) + \sum_{i=1}^{n} K_i X_i \beta \qquad \text{for the Poisson family, and}$$

$$\sum_{i=1}^{n} \{K_i X_i \beta - (\frac{1}{a} + K_i)\ell_n(1 + a\exp(X_i\beta))\} \qquad \text{for the negative binomial family,}$$

where a is an estimator of the dispersion parameter of the negative binomial distribution. These estimators are shown to be strongly consistent and asymptotically normal. GMT (1984b) give the appropriate expression of the asymptotic covariance matrix of the estimated parameters.

GMT (1984a) also present quasi generalised pseudo maximum likelihood methods which need stronger assumptions than pseudo maximum likelihood methods. They show that if the "true" distribution is a member of the LEF, quasi generalised pseudo maximum likelihood estimators (QGPMLE) of the negative binomial regression model (not Poisson model) will be more efficient than the PMLE, provided that the mean and the variance are

well-specified. For the QGPMLE of the negative binomial regression model, the objective function is

$$\sum_{i=1}^{n} \{K_i X_i \beta - (\frac{1}{\hat{\eta}^2} + K_i) \ell n(1 + \hat{\eta}^2 \exp(X_i \beta))\}$$

$$\text{where } \hat{\eta}^2 = \frac{\sum_{i=1}^{n} [(K_i - \exp(X_i \hat{\beta}))^2 - \exp(X_i \hat{\beta})] \exp(2X_i \hat{\beta})}{\sum_{i=1}^{n} \exp(4X_i \hat{\beta})}$$

and $\hat{\beta}$ comes from PMLE. There is no way to be sure that the specified distribution is the "true" one, but, if estimators based on weaker assumptions are close to MLE, on can reasonably assume that the model is not misspecified. QGPMLE of the negative binomial regression model are presented in Table II.8. As one can see, coefficients are very close to those of the MLE of the negative binomial regression model based on strong distributional assumptions.

Gouriéroux and Monfort (1989b) also propose some simulation based methods for models with heterogeneity by replacing mathematical expectations in the objective functions by empirical means from simulated values of the error term. The sampling method is also important for the choice of the estimation method to be used with count data. The sample is said to be exogenous, as in our case, when decision makers (individuals) are sampled and their choices (accidents) observed. In that case of exogenous sampling, maximum likelihood and pseudo maximum likelihood methods are appropriate. The sample is said to be endogenous when choices are sampled and decision makers observed. For example, if one had sampled 50% of individuals with no accident and 50% of individuals with accidents, the resulting sample would overrepresent the population of drivers with accidents since 6.5% of the entire population had at least one accident. In that case of endogenous sampling, one should use some other estimation methods. Statistical methods have been developed to deal with that kind of problem [see Manski-McFadden (1981) for a survey]. Manski-Lerman (1977) propose a weighted maximum likelihood method called Weight Exogenous Sampling Maximum Likelihood (WESML) method. Gouriéroux and Monfort (1989a) showed that Exogenous Sampling Berkson (ESB) method keeps interesting properties when the observations are endogenously stratified and the proposed Modified Exogenous Sampling Berkson (MESB)

method which is shown to be asymptotically more efficient than the ESB method and the WESML methods.

### II.5 Comparison of estimated individual probabilities of accident by the different models

It is difficult to compare the econometric results of the different models insofar as these models are not directly comparable on a statistical basis. For instance, the categorical models of subsection II.2 provide the probability of being in one category or another, that is to be in the category of those with no accident or in the category of those with at least one accident (for the linear probability model, the probit model and the logit model; for the multinomial logit, the latter category is divided into two categories), while the Poisson and negative binomial models, both the homogeneous and heterogeneous drivers forms, give the probability of having 0, 1, 2, 3, etc. accidents. Since the Poisson model is a limited case of the negative binomial model and since the homogeneous drivers Poisson and negative binomial models are special cases of the corresponding heterogeneous drivers models, rigorous hypothesis tests can be performed to discriminate between the models. They were already discussed in the previous section. However, the reader may observe that all models provide the same qualitative results in the sense that the significant variables are essentially the same in the different models. Therefore, any of these models can be estimated when the objective is limited to the identification of significant explanatory variables.

In this section, the econometric results are compared on the basis of the predicted probability of accident or expected number of accidents for given individuals, that is for different X vectors. Tables II.9 to II.14 present the results. Tables II.9 to II.11 give the predicted probabilities of having at least one accident in 1982-1983, given different combinations of the individual's past driving experience variables, and Tables II.12 to II.14 do the same for the predicted expected number of accidents. We observe that the predictions differ with the distributional assumptions. Consequently, in order to obtain consistent quantitative predictions, it is important to select the appropriate model.

## III. THE DETERMINATION OF AN OPTIMAL PRICING FORMULA FOR AUTOMOBILE INSURANCE

In the presence of moral hazard between the principal (insurer) and the agent (insured), the traditional one-period insurance contracts that do not take into account this asymmetrical information problem are not optimal. Many authors have investigated optimal insurance

pricing in single-period models when moral hazard is present [see, for example, Holmstrom (1989) and Shavell (1979)].

More recently, Radner (1981, 1985), Rubinstein and Yaari (1983), Boyer and Dionne (1985, 1987a) have shown that multiperiod contracts are pareto superior to a repetition of single-period contracts when the number of periods is large and the discount factor is close to one. One interesting characteristic of these contracts for automobile insurance concerns the use, by the principal, of statistical information obtained over time from the agent to relate insurance premiums to the individual's past experience. Since the insurer does not observe the individual's driving behavior, he may then use his past experience to enforce more care activities in each period.

The statistical results in the previous sections of this paper clearly show that the individual's past driving experience is a good predictor of risk. In this section, we will show how the Rubinstein-Yaari model for the pricing of insurance in the presence of moral hazard can be applied to the negative binomial model with a regression component.

Considering the Rubinstein-Yaari model, we can reformulate their contract in order to introduce all the information contained in the regression component[1]. The Rubinstein-Yaari "no claims discount" contract is expressed in terms of the observed past average claim and the expected claim obtained from a proper behavior regarding self-protection. In each period, the insurer sets the level of the insurance premium, P(t), as a function of past experience and the insured must decide whether to purchase insurance or not and must choose the level of care activities. The insurer observes the insured's number of accidents when the latter buy insurance. Rubinstein and Yaari propose the following "no-claim discounts" strategy for the leader (insurer) in a Stackelberg game:

$$P(1) = P^*$$

$$P(t) = \begin{bmatrix} P^* & \text{if } \frac{1}{N(t)} \sum_{s=1}^{t-1} L^s < E(L) + \alpha^{N(t)} \\ PP & \text{otherwise} \end{bmatrix}$$

---

1 For a reformulation of the Rubenstein-Yaari contract in terms of probabilities of accidents estimated with a probit model, see Boyer and Dionne (1989).

where P(t) is the premium for period $t > 1$; $P^*$ is the full information premium; $L^s$ is the observed number of claims in period s; N(t) is the number of periods during which the agent was insured; $\alpha^{N(t)}$ is a statistical margin of error, on which we shall return; PP is a penalty premium.

The role of the penalty premium is important. It does not correspond to a simple adjustment to past experience of the previous period premium. Rubinstein and Yaari showed that, in order to have an optimal solution, PP has to be set such that the individual would prefer to self-insure than to buy insurance in a one-period contract. Then, in order to obtain the commitment of the insured to this type of contract, the insurer must use a statistical margin of error that guarantees to those who behave properly that they will not be penalized often.

If we now consider an economy with many types of individuals according to their risk experience and driving skills, the Rubinstein-Yaari optimal multiperiod contract can be extended when individuals' characteristics are observable or can be approximated by a vector of observable variables. To see this, let us rewrite a "Rubinstein and Yaari contract" corresponding to the negative binomial model:

$$P(1) = P_i^*$$

$$P(t) = \begin{cases} P_i^* & \text{if} \quad \frac{1}{N(t)} \sum_{s=1}^{t-1} L^s < \exp(x_i^* \beta) + \alpha_i^{N(t)} \\ PP_i & \text{otherwise} \end{cases}$$

where $x_i^*$ is the vector of characteristics of a type i individual evaluated at the socially-efficient level of self-protection activities generating different values of $D_i^*$, $S_i^*$ and $Z_i^*$.

The above contract can be interpreted as follows: considering that $\exp(x_i^* \beta)$ is the socially-optimal expected number of accidents corresponding to a type i individual, this individual is offered, at period t, a full insurance contract at the $P_i^*$ premium if the average number of past accidents over the (t - 1) previous periods is less than the socially-optimal expected number of accidents plus a statistical margin of error.

To illustrate how such a contract can be designed from our results, we shall consider a driver who corresponds to the average of individual vectors of characteristics ($\bar{x}_i$). We shall simulate "histories" of accidents randomly by a negative binomial distribution with parameters (1.4663, 27.6306) implying that $\exp(\bar{x}_i\beta) = 0.0531$ with a corresponding variance equal to 0.0550. Let us assume that the socially-optimal expected number of accidents is also equal to 0.0531. Assuming that the individual takes insurance in each period, the corresponding statistical margin of error, $\alpha_i^t$, is then equal to

$$\sqrt{\frac{2\gamma(0.0550)\ \text{loglog}t}{t}}$$

where $\gamma$ is a policy parameter and $\alpha_i^t$ is obtained from the law of the iterated logarithm [see Radner (1981) for more details]. In each $t \geq 3$ period, the individual will be charged the $P_i^*$ premium if the past observed frequency of accidents $1/t \sum_{s=1}^{t-1} L^s$ is less than $\exp(\bar{x}_i\beta) + \alpha_i^t$ where $\alpha_i^t$ is defined as above. He will be charged a penalty premium, otherwise. With an appropriately chosen $PP_i$ (penalty premium), the individual is induced to select the proper level of care . However, even with the proper or optimal level of care, the insured may be unlucky and have to pay $PP_i$ often. However, $\alpha_i^t$ prevents the insurer from penalizing too often those who choose the appropriate level of prevention.

The simulation permitted the verification of the above assertion. We simulated 1,000 cases of a 50-period contract and of a 100-period contract and 100 cases of a 1,000-period contract. The results for $\gamma = 1.05$ are summarized in the following table.

Although Rubinstein and Yaari's results were derived from an infinite horizon model without discounting, it appears reasonable to design such a contract spanning a 50-year horizon or a 100-semester horizon. In Quebec, for example, automobile insurance for bodily injuries is offered by a public monopoly and drivers must buy insurance in order to drive. Then their average length of contracts is about 50 years or 100 semesters. In both cases, they would be penalized less than 8 % of the total number of periods if they behave properly. However, an individual who does not produce the optimal level of care and generates an accident frequency of 0.07 instead of 0.053, for example, will be penalized for more than 15 % of the periods.

**Average Number (#) and Percentage (%) of Periods for which the Average Individual Will Be Wrongly Penalized**

| Contract | | Periods with $PP_i$ | |
|---|---|---|---|
| | | (#) | (%) |
| 50-period | (1,000 cases) | 4.17 | 8.34 |
| 100-period | (1,000 cases) | 7.03 | 7.03 |
| 1,000-period | (100 cases) | 27.80 | 2.78 |

## CONCLUSION

In this study, we reviewed and estimated eight econometric models of accident distributions. The estimation of the appropriate statistical model is important to implement insurance contracts that reduce the inefficiencies associated to asymmetrical information. We used a set of data on 19,013 car drivers for which we had five types of information: the personal characteristics, the number of accidents, the number of demerit points, the number of license suspensions for criminal offenses and the number of days the driving permit was valid during each period. We have estimated four categorical models (linear probability, probit, logit and multinomial logit) and four count data models (Poisson and negative binomial models with and without individual characteristics in the regression component).

It is difficult to compare the econometric results of the different models since some of these models are not nested. However, it was shown that the negative binomial model with a regression component produces a reasonable approximation of the true distribution of accidents. Different statistical tests rejected the Poisson models (with and without a regression component) and the negative binomial model without individual characteristics. It was also observed that all estimated models provide the same qualitative results (essentially the same significant variables), but differ when predictions of either the probabilities of accidents or the expected number of accidents were made. For quantitative predictions, it is important to select the appropriate model. Finally, it was shown that, in all models, the individual's past driving experience is a good predictor of risk.

In section III, we used the statistical results obtained from the negative binomial model with a regression component to show how the Rubinstein-Yaari model of insurance tarification

in the presence of moral hazard could be applied. Other types of tarification models could also be investigated.

## REFERENCES


Belsey, David A., E. Kuh and R.E. Welsch (1980), *Regression Diagnostics, Identifying Influential Data and Sources of Multicollinearity*, New York, Wiley.

Boyer, M. and G. Dionne (1985), "Sécurité routière: responsabilité pour négligence et tarification", *Canadian Journal of Economics/Revue canadienne d'économique* 18, November, 814-830.

Boyer, M. and G. Dionne (1987a), "The Economics of Road Safety", *Transportation Research B* 21(5), 413-431.

Boyer, M. and G. Dionne (1987b), "Description and Analysis of the Québec Automobile Insurance Plan", *Canadian Public Policy/Analyse de politiques* XIII(2), 181-195.

Boyer, M. and G. Dionne (1989), "An Empirical Analysis of Moral Hazard and Experience Rating", *Review of Economics and Statistics* LXXI(1), 128-134.

Boyer, M., G. Dionne and C. Vanasse (1988), "Infractions au code de la sécurité routière, infractions au code criminel et accidents automobiles", R.A.A.Q. report and Discussion Paper No. 583, Centre de recherche sur les transports, 85 pages.

Cameron, A.C. and P.K. Trivedi (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimations and Tests", *Journal of Applied Econometrics* 1, 29-53.

Cameron, A.C., P.K. Trivedi, F. Milne and J. Piggott (1988), "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia", *Review of Economic Studies* 55, 85-106.

Dahlby, B.A. (1983), "Adverse Selection and Statistical Discrimination", *Journal of Public Economics* 20, 121-131.

Ferreira, J. (1974), "The Long-Term Effects of Merit-Rating Plans on Individual Motorists", *Operations Research* 22, September-October, 954-978.

Gilbert, C.L. (1979), "Econometric Models for Discrete Economic Processes", Discussion Paper, University of Oxford.

Gouriéroux, C. and A. Montfort (1989a), "Econometrics Based on Endogenous Samples", INSEE, Discussion Paper 89.3.

Gouriéroux, C. and A. Montfort (1989b), "Simulation Based Inference in Models with Heterogeneity", INSEE, Discussion Paper 89.5.

Gouriéroux, C. and A. Montfort (1989c), "Econometrics of Count Data: The A.L.D.P. Model", paper presented at the European Society Econometric Meeting, Munich, September, 50 pages.

Gouriéroux, C., A. Montfort and A. Trognon (1984a), "Pseudo Maximum Likelihood Methods: Theory", *Econometrica* 52, 681-700.

Gouriéroux, C., A. Montfort and A. Trognon (1984b), "Pseudo Maximum Likelihood Methods: Applications to Poisson Models", *Econometrica* 52, 701-720.

Greenwood, M. and G.U. Yule (1920), "An Inquiry into the Nature of Frequency Distribution of Multiple Happenings", *Journal of the Royal Statistical Society* A(83), 255-279.

Hausman, J., B.H. Hall and Z. Griliches (1984), "Econometric Models for Count Data with an Application to the Patents-R & D Relationship", *Econometrica* 52, 909-938.

Heckman, J. and G. Borjas (1980), "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model for Heterogeneity and State Dependence", *Econometrica* 47, 247-283.

Holland, P.W. (1973), "Poisson Regression", Discussion Paper, Princeton University.

Holmstrom, B. (1979), "Moral Hazard and Observability", *Bell Journal of Economics* 10, Spring, 74-91.

Johnson, P.D. and G.B. Hey (1971), "Statistical Studies in Motor Insurance", *The Journal of the Institute of Actuaries* 97, 199-232.

Lemaire, J. (1985), *Automobile Insurance*: *Actuarial Models*, Boston, Kluwer-Nighoff.

Maddala, G.S. (1977), *Econometrics*, New York, McGraw-Hill.

Manski, C.F. and S.R. Lerman (1977), "The Estimation of Choice Probabilities from Choice Based Samples", *Econometrica* 45, 1977-1988.

Manski, C.F. and D. McFadden (1981), "Alternative Estimators and Sample Designs for Discrete Choice Analysis", in *Structural Analysis of Discrete Data with Econometric Applications*, C.F. Manski and D. McFadden (eds.), M.I.T. Press.

Patil, G.P. (1970), *Random Counts in Models and Structures*: *Volume 1*, The Pennsylvania State University Press, Pennsylvania.

Radner, R. (1981), "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship", *Econometrica* 49, September, 1127-1148.

Radner, R. (1985), "Repeated Principal-Agent Games with Discounting", *Econometrica* 53, September, 1173-1199.

Rose, N. (1988), "Profitability and Product Quality: Financial Indicators and Airline Safety Performance", mimeo, M.I.T.

Rubinstein, A. and M.E. Yaari (1983), "Repeated Insurance Contracts and Moral Hazard", *Journal of Economic Theory* 30, June, 79-97.

Seal, H.L. (1969), *Stochastic Theory of a Risk Business*, New York, Wiley.

Shavell, S. (1979), "Risk Sharing and Incentives in the Principal and Agent Relationship", *Bell Journal of Economics* 10, Spring, 55-74.

Theil, H. (1971), *Principles of Econometrics*, New York, Wiley.

**Table A.1**
**List of Variables Used**

---

AGE: 8 dichotomous variables

A16 = 1 if, on August 1, 1982, the driver is under 16 years of age (some drivers fall into this category since the sample was taken on May 16, 1983), 0 otherwise

A1619 = reference group

A2024 = 1 if, on August 1, 1982, the driver is between 20 and 24, 0 otherwise,

etc.

SEX: two dichotomous variables

SEXF = reference group

SEXM = 1 if male, 0 otherwise

EXPERIENCE: seven dichotomous variables to measure the number of years since a driver obtained a first permit

EXP0 = 1 if permit obtained in 1982–83, 0 otherwise

EXPM1= 1 if, on August 1, 1982, the permit has been held for less than a year, 0 otherwise

EXP1 = 1 if permit was obtained between one and two years ago, 0 otherwise (reference group)

etc.

PLACE OF RESIDENCE: eleven dichotomous variables to capture the effect of the immediate environment of the individual driver

REG1 = 1 if the driver lives in the lower St. Lawrence or Gaspé regions, 0 otherwise

REG2 = Saguenay–Lac St.Jean

REG3 = Québec city region

REG4 = Trois–Rivières
REG5 = Eastern Townships
REG6 = Montréal region: reference group
REG7 = Laurentians–Lanaudière
REG8 = Montérégie
REG9 = Outaouais
REG10 = Abitibi–Témiscamingue
REG11 = North–Shore and Nouveau–Québec

DRIVING RESTRICTIONS: 13 dichotomous variables to capture the effect of driving restrictions appearing on the driver license (type of license)

RTSA = 1 if the driver must wear glasses, 0 otherwise
RTSB = can drive during daylight only
RTSCG = must wear a hearing aid (C) or must drive a car with power brakes (L, N, P) or must drive a car with a left accelerator (R) or is subject to medical restrictions (S) or other restrictions (T) or has a temporary permit (W).
RTSD = must get a medical exam regularly
RTSH = can only drive a vehicle whose total weight is less than 5500 kg
RTSJ = must drive an automatic transmission equipped car
RTSK = must drive a car with power steering
RTSM = cannot drive a business vehicle, a bus, a minibus or a taxi
RTSO = cannot drive an emergency vehicle
RTSQ = must drive a car equipped with manual signals or a magic eye
RTSU = has a license valid for 6 months only
RTSY = cannot drive a taxi or an ambulance
RTSØ = no restrictions

CLASS OF DRIVING LICENSE: 9 dichotomous variables to isolate the effect of the class of driving license (since class 41 and class 42 are almost mutually exclusive, only class 42 is used as a variable; since no driver in our sample has class 61 only, this class is not used as a variable).

CL1112 = 1 if driver can drive a bus (includes 13,41,42,61), 0 otherwise
CL13 = minibus (includes 41,42,61)
CL21 = vehicle or set of vehicles whose weight may exceed 11000 kg (includes 22, 41, 42, 61)

CL22 = vehicle whose weight may exceed 11000 kg (includes 41, 42,61)

CL31 = taxis (includes 41, 42, 61)

CL41 = vehicles or set of vehicles whose weight is less than 11000 kg (includes 42,61)

CL42 = vehicles of class 41 except public vehicles (includes 61)

CL54 = motorcycle (includes 55, 56, 61)

CL55 = motorcycle of less than 400 $cm^3$ (includes 56, 61)

CL56 = velomotor (includes 61)

CL61 = tractor

VALIDITY: three continuous variables to measure the number of days for which the driving license was valid in a given year

VALA : number of days during which the driver's license was valid during 1980–81

VALB : number of days during which the driver's license was valid during 1981–82

VALC : number of days during which the driver's license was valid during 1982–83

DEMERIT POINTS: variable D is the number of demerit points cumulated over the two–year period between August 1980 and July 1982 (see table below)

SUSPENSIONS OF LICENSE: variable S is the number of license suspensions for criminal offenses (see table below) in effect during the year 1981–82

PAST INVOLVEMENTS IN ACCIDENTS: variable Z is the number of accidents in which the driver was involved over the two–year period between August 1980 and July 1982

**Table A.2**

**Partial List of Driving Infractions and Corresponding Number of Demerit Points (D)**

(bringing a license suspension if 12 or more points are cumulated over a two–year period; the license suspension is for 3, 6 and 12 months respectively for the first, second and third offense)

| INFRACTIONS | # OF POINTS |
|---|---|
| – Illegal backward movement | 2 |
| – Not stopping at a stop sign | 2 |
| – Not stopping at a red light | 3 |
| – Illegal passing: *on left | 2 |
| *on right | 1 |
| – Illegal crossing of a continuous line | 2 |
| – Driving on the passage reserved for cars in the other direction | 4 |
| – Racing | 6 |
| – Not stopping at Railway Crossings while driving a bus or a dangerous material carrier | 9 |
| – Not stopping for a school bus with blinking lights on | 9 |
| – Excess speed: | |
| * 1 to 14 km/h over limit | 1 |
| *15 to 29 km/h over limit | 2 |
| *30 to 44 km/h over limit | 3 |
| *45 + km/h over limit | 4 + |

**Table A.3**

**Partial List of Criminal Offenses (S)**

(bringing a license suspension of 3, 6 and 12 months respectively for the first, second and third offense over a two–year period)

- Criminal negligence causing death or bodily injuries
- Hit and run
- Dangerous driving
- Refusal of a breathalyzer test
- Driving (actual or intended) under the influence of alcohol (80 mg/100 ml of blood)
- Convicted of manslaughter in a car accident

Table I.1*

**Drivers According to Their Number of Demerit Points Accumulated in 1980–1981 and 1981–1982 and Their Accidents in 1982–1983**

| Number of Demerit Points Accumulated in 1980–1981 and 1981–1982 | Number of Accidents in 1982–1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| 0 | 14,074<br>95.1<br>79.1 | 681<br>4.6<br>59.8 | 36<br>0.2<br>45.6 | 6<br>0<br>54.5 | 14,797<br>100<br>77.8 |
| 1 to 5 | 3,046<br>89.4<br>17.1 | 332<br>9.7<br>29.1 | 28<br>0.8<br>35.4 | 2<br>0.1<br>18.2 | 3,408<br>100<br>17.9 |
| 6 to 11 | 575<br>83.3<br>3.2 | 104<br>15.1<br>9.1 | 9<br>1.3<br>11.4 | 2<br>0.3<br>18.2 | 690<br>100<br>3.6 |
| 12 and more | 89<br>75.4<br>0.5 | 22<br>18.6<br>1.9 | 6<br>5.1<br>7.6 | 1<br>0.8<br>9.1 | 118<br>100<br>0.6 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

**Table I.2***

**Drivers According to Their Number of Accidents in 1980–1981 and 1981–1982 and Their Accidents in 1982–1983**

| Number of Accidents in 1980–1981 and 1981–1982 | Number of Accidents in 1982–1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| 0 | 15,449<br>94.5<br>86.9 | 842<br>5.2<br>73.9 | 54<br>0.3<br>68.4 | 4<br>0<br>36.4 | 16,349<br>100<br>86 |
| 1 | 1,980<br>89.5<br>11.1 | 215<br>9.7<br>18.9 | 16<br>0.7<br>20.3 | 2<br>0.1<br>18.2 | 2,213<br>100<br>11.6 |
| 2 | 296<br>81.3<br>1.7 | 59<br>16.2<br>5.2 | 6<br>1.6<br>7.6 | 3<br>0.8<br>27.3 | 364<br>100<br>1.9 |
| 3 and more | 59<br>67.8<br>0.3 | 23<br>26.4<br>2.0 | 3<br>3.4<br>3.8 | 2<br>2.3<br>18.2 | 87<br>100<br>0.5 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

Table I.3*

Drivers According to Their Number of Suspensions for Criminal Code Offenses in 1980–1981 and 1981–1982 and Their Accidents in 1982–1983

| Number of Suspensions in 1981–1982 | Number of Accidents in 1982–1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| 0 | 17,691<br>93.6<br>99.5 | 1,119<br>5.9<br>98.2 | 76<br>0.4<br>96.2 | 11<br>0.1<br>100 | 18,897<br>100<br>99.4 |
| 1 | 87<br>80.6<br>0.5 | 19<br>17.6<br>1.7 | 2<br>1.9<br>2.5 | 0<br>0<br>0 | 108<br>100<br>0.6 |
| 2 and more | 6<br>75<br>0 | 1<br>12.5<br>0.1 | 1<br>12.5<br>1.3 | 0<br>0<br>0 | 8<br>100<br>0 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

**Table I.4***

**Drivers According to Their Sex and Their Accidents in 1982–1983**

| Sex | Number of Accidents in 1982–1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| Male | 10,297<br>91.3<br>57.9 | 902<br>8<br>79.2 | 71<br>0.6<br>89.9 | 10<br>0.1<br>90.9 | 11,280<br>100<br>59.3 |
| Female | 7,487<br>96.8<br>42.1 | 237<br>3.1<br>20.8 | 8<br>0.1<br>10.1 | 1<br>0<br>9.1 | 7,733<br>100<br>40.7 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

**Table I.5***

**Drivers According to Their Age Group and Their Accidents in 1982-1983**

| Age Groups | Number of Accidents in 1982-1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| 15 and less | 53<br>96.4<br>0.3 | 2<br>3.6<br>0.2 | 0<br>0<br>0 | 0<br>0<br>0 | 55<br>100<br>0.3 |
| 16-19 | 1,209<br>90.4<br>6.8 | 113<br>8.5<br>9.9 | 14<br>1<br>17.7 | 1<br>0.1<br>9.1 | 1,337<br>100<br>7.0 |
| 20-24 | 2,478<br>90.6<br>13.9 | 231<br>8.4<br>20.3 | 23<br>0.8<br>29.1 | 3<br>0.1<br>27.3 | 2,735<br>100<br>14.4 |
| 25-34 | 4,981<br>93.6<br>28.0 | 320<br>6.0<br>28.1 | 19<br>0.4<br>24.1 | 4<br>0.1<br>36.4 | 5,324<br>100<br>28.0 |
| 35-44 | 3,721<br>94.3<br>20.9 | 211<br>5.3<br>18.5 | 11<br>0.3<br>13.9 | 2<br>0.1<br>18.2 | 3,945<br>100<br>20.7 |
| 45-54 | 2,530<br>94.4<br>14.2 | 145<br>5.4<br>12.7 | 6<br>0.2<br>7.6 | 0<br>0<br>0 | 2,681<br>100<br>14.1 |
| 55-64 | 1,791<br>95.7<br>10.1 | 77<br>4.1<br>6.8 | 3<br>0.2<br>3.8 | 1<br>0<br>9.1 | 1,872<br>100<br>9.8 |
| 65 and over | 1,021<br>96.0<br>5.7 | 40<br>3.8<br>3.5 | 3<br>0.3<br>3.8 | 0<br>0<br>0 | 1,064<br>100<br>5.6 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

**Table I.6***
**Drivers According to Their Class of Permit and Their Accidents in 1982–1983**

| Class of Permit (see Table A.1 for exact definitions) | Number of Accidents in 1982–1983 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 and More | TOTAL |
| 11, 12 (bus driver) | 201<br>85.2<br>1.1 | 28<br>11.9<br>2.5 | 4<br>1.7<br>5.1 | 3<br>1.3<br>27.3 | 236<br>100<br>1.2 |
| 13 (mini bus) | 20<br>95.2<br>0.1 | 1<br>4.8<br>0.1 | 0<br>0<br>0 | 0<br>0<br>0 | 21<br>100<br>0.1 |
| 21 (trailer trucks) | 3,694<br>90.8<br>20.8 | 342<br>8.4<br>30 | 26<br>0.6<br>32.9 | 5<br>0.1<br>45.5 | 4,067<br>100<br>21.4 |
| 22 (trucks) | 43<br>76.8<br>0.2 | 11<br>19.6<br>1.0 | 2<br>3.6<br>2.5 | 0<br>0<br>0 | 56<br>100<br>0.3 |
| 31 (taxis) | 107<br>79.3<br>0.6 | 24<br>17.8<br>2.1 | 3<br>2.2<br>3.8 | 1<br>0.7<br>9.1 | 135<br>100<br>0.7 |
| 41 (cars and small trucks) | 6,087<br>91.6<br>34.2 | 507<br>7.6<br>44.5 | 42<br>0.6<br>53.2 | 1<br>0.1<br>63.6 | 6,643<br>100<br>34.9 |
| 42 (41; private vehicles only) | 11,694<br>94.6<br>65.8 | 632<br>5.1<br>55.5 | 37<br>0.3<br>46.8 | 4<br>0<br>36.4 | 12,367<br>100<br>65.0 |
| 54 (motorcycles) | 14,426<br>93.8<br>81.8 | 897<br>5.8<br>78.8 | 55<br>0.8<br>69.6 | 9<br>0.1<br>81.8 | 15,387<br>100<br>80.9 |
| 55 (small motorcycles) | 31<br>93.9<br>0.2 | 1<br>3<br>0.1 | 1<br>3<br>1.3 | 0<br>0<br>0 | 33<br>100<br>0.2 |
| 56 (velomotor) | 8<br>88.9<br>0 | 1<br>11.1<br>0.1 | 0<br>0<br>0 | 0<br>0<br>0 | 9<br>100<br>0 |
| TOTAL | 17,784<br>93.5<br>100 | 1,139<br>6<br>100 | 79<br>0.4<br>100 | 11<br>0.1<br>100 | 19,013<br>100<br>100 |

* Each cell contains three numbers. The first one indicates the number of drivers in the cell, the second expresses the number as a percentage of the row total and the third expresses it as a percentage of the column total.

**Table II.1**

**Estimates of the Univariate Poisson and Negative Binomial Distributions**

| Number of Accidents k | Observed Number of Individuals with k Accidents in 1982–1983 $n_k$ | Predicted Number of Individuals with k Accidents in 1982–1983 | |
|---|---|---|---|
| | | Poisson ($\hat{\lambda}$=0.0701) | Negative Binomial ($\hat{a}$=0.6960; $\hat{\tau}$=9.9359) |
| 0 | 17,784 | 17,726.60 | 17,785.28 |
| 1 | 1,139 | 1,241.86 | 1,132.05 |
| 2 | 79 | 43.50 | 88.79 |
| 3 | 9 | 1.02 | 7.1 |
| 4 | 2 | 0.02 | 0.61 |
| 5+ | 0 | 0 | 0 |
| | 19,013 | $\chi^2$ = 133.06 | $\chi^2$ = 2.21 |
| | | $\chi^2_{2,95}$ = 5.99 | $\chi^2_{1,95}$ = 7.82 |
| | | LL = -4,950.28 | LL = -4,916.78 |

**Table II.2**

**Least Squares Estimates**

| Variable | Model 1 | | Model 2 | |
|---|---|---|---|---|
| D | — | | .011 | (12.37) ** |
| S | — | | .057 | ( 2.77) ** |
| Z | .048 | (12.48) ** | .040 | (10.05) ** |
| CONSTANT | .049 | ( 2.33) ** | .046 | ( 2.19) ** |
| A16 | -.039 | (-1.11) | -.036 | (-1.05) |
| A2024 | -.004 | (-0.42) | -.007 | (-0.65) |
| A2534 | -.033 | (-3.08) ** | -.030 | (-2.82) ** |
| A3544 | -.044 | (-3.91) ** | -.038 | (-3.40) ** |
| A4554 | -.046 | (-3.88) ** | -.037 | (-3.14) ** |
| A5564 | -.058 | (-4.74) ** | -.047 | (-3.88) ** |
| A65+ | -.058 | (-4.21) ** | -.044 | (-3.22) ** |
| SEXM | .044 | (10.74) ** | .034 | ( 8.23) ** |
| EXP0 | -.015 | (-0.75) | -.008 | (-0.39) |
| EXPM1 | -.001 | (-0.07) | .3E-3 | ( 0.02) |
| EXP2 | -.018 | (-1.19) | -.016 | (-1.06) |
| EXP35 | -.028 | (-1.78) * | -.028 | (-1.75) * |
| EXP610 | -.030 | (-1.77) * | -.031 | (-1.84) * |
| EXP11 | -.025 | (-1.44) | -.026 | (-1.49) |
| REG1 | .008 | ( 0.81) | .010 | ( 1.03) |
| REG2 | -.3E-3 | (-0.03) | .004 | ( 0.43) |
| REG3 | .004 | ( 0.69) | .007 | ( 1.30) |
| REG4 | .011 | ( 1.44) | .011 | ( 1.56) |
| REG5 | .5E-3 | ( 0.05) | .001 | ( 0.12) |
| REG7 | .008 | ( 1.26) | .009 | ( 1.32) |
| REG8 | -.6E-4 | (-0.01) | .3E-4 | ( 0.01) |
| REG9 | .028 | ( 3.08) ** | .031 | ( 3.45) ** |
| REG10 | .017 | ( 1.49) | .015 | ( 1.32) |
| REG11 | .016 | ( 1.18) | .017 | ( 1.30) |
| RTSA | -.035 | (-2.75) ** | -.033 | (-2.66) ** |
| RTSB | .014 | ( 0.49) | .015 | ( 0.55) |
| RTSCG | -.021 | (-1.32) | -.020 | (-1.24) |
| RTSD | .1E-3 | ( 0.01) | .5E-3 | ( 0.03) |
| RTSH | .018 | ( 1.36) | .017 | ( 1.36) |
| RTSJ | .096 | ( 1.69) * | .095 | ( 1.68) * |
| RTSK | -.060 | (-0.97) | -.061 | (-0.98) |
| RTSM | -.027 | (-1.46) | -.026 | (-1.44) |
| RTSO | .014 | ( 0.48) | .017 | ( 0.61) |
| RTSQ | -.071 | (-0.97) | -.066 | (-0.90) |
| RTSU | .058 | ( 2.14) ** | .060 | ( 2.23) ** |
| RTSY | -.077 | (-2.66) ** | -.075 | (-2.59) ** |
| RTSØ | -.026 | (-2.02) ** | -.026 | (-2.03) ** |

**Table II.2 (cont'd)**

| Variable | Model 1 | | Model 2 | |
|---|---|---|---|---|
| CL1112 | .034 | ( 2.08) ** | .035 | ( 2.10) ** |
| CL13 | –.035 | (–0.67) | –.031 | (–0.59) |
| CL21 | .015 | ( 1.78) * | .016 | ( 1.89) * |
| CL22 | .135 | ( 4.06) ** | .127 | ( 3.85) ** |
| CL31 | .097 | ( 4.55) ** | .082 | ( 3.88) ** |
| CL42 | .005 | ( 0.73) | .006 | ( 0.88) |
| CL54 | .007 | ( 0.74) | .007 | ( 0.68) |
| CL55 | –.067 | (–1.57) | –.080 | (–1.89) * |
| CL56 | –.001 | (–0.01) | .004 | ( 0.04) |
| VALA | .8E–4 | ( 1.43) | .4E–4 | ( 0.78) |
| VALB | –.1E–3 | (–2.36) ** | –.7E–4 | (–1.59) |
| VALC | .2E–3 | ( 5.24) ** | .2E–3 | ( 4.82) ** |
| $\bar{R}^2$ | 0.0325 | | 0.0407 | |
| F | 13.78 | | 16.52 | |
| nb of variables | 51 | | 53 | |
| nb of observations | 19013 | | 19013 | |

*significant at 90% **significant at 95%

F statistic (to test that $\beta_D = \beta_S = 0$) = 82.06 > $F^{99}_{(2,\infty)}$ = 4.60.

**Table II.3**

**Probit Estimates**

| Variable | MODEL 1 | | | MODEL 2 | | |
|---|---|---|---|---|---|---|
| | Original Coefficient | (t) | Transformed Coefficient | Original Coefficient | (t) | Transformed Coefficient |
| D | - | - | - | 0.055 | ( 9.62)** | 0-1 : 0.006<br>1-5 : 0.028<br>5-31: 0.434 |
| S | - | - | - | 0.290 | ( 2.23)** | 0-1 : 0.039<br>1-2 : 0.057<br>2-3 : 0.077 |
| Z | .257 | (10.12)** | 0-1 : 0.032<br>1-2 : 0.046<br>2-6 : 0.327 | 0.211 | ( 8.11)** | 0-1 : 0.025<br>1-2 : 0.034<br>2-6 : 0.239 |
| CONSTANT | -2.128 | (-10.21)** | NC | -2.295 | (-10.30)** | NC |
| A16 | -0.323 | (-0.95) | -0.027 | -0.304 | (-0.89) | -0.25 |
| A2024 | -0.014 | (-0.18) | -0.002 | -0.028 | (-0.35) | -0.003 |
| A2534 | -0.224 | (-2.60)** | -0.022 | -0.207 | (-2.40)** | -0.020 |
| A3544 | -0.330 | (-3.53)** | -0.031 | -0.292 | (-3.11)** | -0.027 |
| A4554 | -0.350 | (-3.58)** | -0.031 | -0.288 | (-2.93)** | -0.026 |
| A5564 | -0.482 | (-4.62)** | -0.038 | -0.409 | (-3.91)** | -0.033 |
| A65 | -0.468 | (-3.93)** | -0.036 | -0.372 | (-3.11)** | -0.030 |
| SEXM | 0.432 | (11.69)** | 0.044 | 0.370 | ( 9.80)** | 0.037 |
| EXP0 | 0.054 | ( 0.27) | 0.006 | 0.227 | ( 1.08) | 0.029 |
| EXPM1 | 0.050 | ( 0.32) | 0.006 | 0.116 | ( 0.72) | 0.014 |
| EXP2 | -0.157 | (-1.22) | -0.015 | -0.145 | (-1.10) | -0.014 |
| EXP35 | -0.241 | (-1.81)* | -0.023 | -0.245 | (-1.79)* | -0.022 |
| EXP610 | -0.239 | (-1.66)* | -0.023 | -0.251 | (-1.70)* | -0.024 |
| EXP11 | -0.187 | (-1.27) | -0.020 | -0.194 | (-1.29) | -0.021 |
| REG1 | 0.090 | ( 1.10) | 0.010 | 0.107 | ( 1.30) | 0.012 |
| REG2 | 0.004 | ( 0.06) | 0.001 | 0.035 | ( 0.46) | 0.004 |
| REG3 | 0.046 | ( 0.98) | 0.005 | 0.073 | ( 1.55) | 0.008 |
| REG4 | 0.097 | ( 1.61) | 0.011 | 0.106 | ( 1.75)* | 0.012 |
| REG5 | 0.021 | ( 0.25) | 0.002 | 0.025 | ( 0.30) | 0.003 |
| REG7 | 0.072 | ( 1.31) | 0.008 | 0.073 | (1.32) | 0.008 |
| REG8 | 0.008 | ( 0.16) | 0.001 | 0.008 | ( 0.18) | 0.001 |
| REG9 | 0.216 | ( 3.07)** | 0.027 | 0.245 | ( 3.46)** | 0.031 |
| REG10 | 0.142 | ( 1.61) | 0.017 | 0.128 | ( 1.43) | 0.015 |
| REG11 | 0.144 | ( 1.34) | 0.017 | 0.158 | ( 1.46) | 0.019 |

**Table II.3 (cont'd)**

| Variable | MODEL 1 | | | MODEL 2 | | |
|---|---|---|---|---|---|---|
| | Original Coefficient | (t) | Transformed Coefficient | Original Coefficient | (t) | Transformed Coefficient |
| RTSA | -0.266 | (-2.52)** | -0.26 | -0.254 | (-2.40)** | -0.025 |
| RTSB | 0.138 | ( 0.54) | 0.017 | 0.155 | ( 0.60) | 0.019 |
| RTSCG | -0.158 | (-1.05) | -0.015 | -0.146 | (-0.97) | -0.014 |
| RTSD | 0.040 | ( 0.33) | 0.005 | 0.034 | (0.28) | 0.004 |
| RTSH | 0.112 | ( 1.00) | 0.013 | 0.118 | ( 1.04) | 0.014 |
| RTSJ | 0.745 | ( 1.77)* | 0.138 | 0.736 | ( 1.73)* | 0.134 |
| RTSK | -0.533 | (-0.99) | -0.037 | -0.535 | (-0.99) | -0.037 |
| RTSM | -0.272 | (-1.56) | -0.024 | -0.273 | (-1.55) | -0.023 |
| RTSO | 0.104 | ( 0.47) | 0.012 | 0.145 | ( 0.67) | 0.017 |
| RTSQ | -0.650 | (-0.97) | -0.041 | -0.591 | (-0.88) | -0.039 |
| RTSU | 0.343 | ( 1.75)* | 0.048 | 0.366 | ( 1.86)* | 0.52 |
| RTSY | -0.997 | (-2.34)** | -0.49 | -0.977 | (-2.31)** | -0.48 |
| RTSØ | -0.189 | (-1.75)* | -0.22 | -0.186 | (-1.71)* | -0.21 |
| CL11,12 | 0.145 | ( 1.30) | 0.018 | 0.158 | ( 1.41) | 0.019 |
| CL13 | -0.298 | (-0.59) | -0.025 | -0.256 | (-0.51) | -0.022 |
| CL21 | 0.121 | ( 1.79)* | 0.014 | 0.127 | ( 1.88)* | 0.014 |
| CL22 | 0.581 | ( 2.95)** | 0.097 | 0.560 | ( 2.81)** | 0.091 |
| CL31 | 0.438 | ( 3.40)** | 0.066 | 0.359 | ( 2.74)** | 0.050 |
| CL42 | 0.039 | ( 0.68) | 0.004 | 0.045 | ( 0.77) | 0.005 |
| CL54 | 0.038 | ( 0.49) | 0.004 | 0.041 | ( 0.52) | 0.004 |
| CL55 | -0.476 | (-1.32) | -0.035 | -0.579 | (-1.55) | -0.038 |
| CL56 | -0.120 | (-0.20) | -0.012 | -0.104 | (-0.17) | -0.010 |
| VALA | 0.001 | ( 1.49) | NC | 0.001 | ( 0.97) | NC |
| VALB | -0.001 | (-1.75)* | NC | -0.000 | (-0.42) | NC |
| VALC | 0.002 | ( 5.87)** | NC | 0.002 | ( 5.70)** | NC |
| nb of observations | | 19013 | | | 19013 | |
| nb of variables | | 51 | | | 53 | |
| Likelihood ratio | | 620.99 | | | 716.64 | |
| Mean est'd prob. of accident | | 0.065 | | | 0.065 | |
| Est'd prob. of accident (avg. ind.) | | 0.053 | | | 0.052 | |
| (Standard error) | | (0.002) | | | (0.002) | |

(* = significant at 90%; ** = significant at 95%; NC = not calculated)

Log likelihood ratio test for the restrictions $\beta_D = \beta_S = 0$: $-2(LL_1 - LL_2) = 95.40 > \chi^2_{2,99} = 9.21$.

**Table II.4**

**Logit Estimates**

| Variable | Coefficient | (t) | $+\pi/\sqrt{3}$ |
|---|---|---|---|
| D | 0.1007 | (9.62)** | 0.0555 |
| S | 0.5454 | (2.30)** | 0.3007 |
| Z | 0.3921 | (8.15)** | 0.2162 |
| CONSTANT | -4.3752 | (-9.32)** | -2.4122 |
| A16- | -0.5871 | (-0.78) | -0.3237 |
| A2024 | -0.0363 | (-0.23) | -0.0200 |
| A2534 | -0.4116 | (-2.38)** | -0.2269 |
| A3544 | -0.5932 | (-3.13)** | -0.3270 |
| A4554 | -0.5846 | (-2.94)** | -0.3223 |
| A5564 | -0.8378 | (-3.90)** | -0.4619 |
| A65+ | -0.8062 | (-3.22)** | -0.4445 |
| SEXM | 0.7960 | (9.74)** | 0.4389 |
| EXP0 | 0.4945 | (1.14) | 0.2726 |
| EXPM1 | 0.2326 | (0.71) | 0.1282 |
| EXP2 | -0.3317 | (-1.28) | -0.1829 |
| EXP35 | -0.5049 | (-1.87)* | -0.2784 |
| EXP610 | -0.5190 | (-1.78)* | -0.2861 |
| EXP11 | -0.3914 | (-1.31) | -0.2158 |
| REG1 | 0.1933 | (1.14) | 0.1066 |
| REG2 | 0.0737 | (0.47) | 0.0406 |
| REG3 | 0.1419 | (1.46) | 0.0782 |
| REG4 | 0.2066 | (1.68)* | 0.1139 |
| REG5 | 0.0116 | (0.07) | 0.0064 |
| REG7 | 0.1473 | (1.30) | 0.0812 |
| REG8 | 0.0034 | (0.04) | 0.0019 |
| REG9 | 0.5087 | (3.65)** | 0.2805 |
| REG10 | 0.2252 | (1.26) | 0.1242 |
| REG11 | 0.2934 | (1.35) | 0.1618 |
| RTSA | -0.4937 | (-2.30)** | -0.2722 |
| RTSB | 0.3162 | (0.58) | 0.1743 |
| RTSGC | -0.3042 | (-0.97) | -0.1677 |
| RTSD | 0.0932 | (0.36) | 0.0514 |
| RTSH | 0.2566 | (1.13) | 0.1415 |
| RTSJ | 1.4925 | (1.92)* | 0.8229 |
| RTSK | -1.0066 | (-0.10) | -0.5550 |
| RTSM | -0.5311 | (-1.43) | -0.2928 |
| RTSO | 0.2887 | (0.67) | 0.1592 |
| RTSQ | -1.2254 | (-0.94) | -0.6756 |

**Table II.4 (cont'd)**

| Variable | Coefficient | (T-Ratio) | $+\pi/\sqrt{3}$ |
|---|---|---|---|
| RTSU | 0.6966 | (1.84)* | 0.3841 |
| RTSY | -2.0983 | (-2.12)** | -1.1569 |
| RTSØ | -0.3395 | (-1.54)* | -0.1872 |
| CL1112 | 0.2748 | (1.31) | 0.1515 |
| CL13 | -0.4290 | (-0.41) | -0.2365 |
| CL21 | 0.2822 | (2.09)* | 0.1556 |
| CL22 | 1.0336 | (2.95)** | 0.5699 |
| CL31 | 0.6556 | (2.81)** | 0.3615 |
| CL42 | 0.0914 | (0.79) | 0.0504 |
| CL54 | 0.0601 | (0.39) | 0.0331 |
| CL55 | -1.1047 | (-1.46) | -0.6091 |
| CL56 | -0.0998 | (-0.09) | -0.0550 |
| VALA | 0.9721E-3 | (0.87) | 0.5359E-3 |
| VALB | -0.5243E-4 | (-0.06) | -0.2891E-4 |
| VALC | 0.4193E-2 | (5.50)** | 0.2312E-2 |

| | |
|---|---|
| LL | -4201.0 |
| Nb of parameters | 53 |
| Nb of observations | 19013 |

* = significant at 90%; ** = significant at 95%

Log likelihood ratio test for the restrictions $\beta_D = \beta_S = 0$:
$-2(LL_1 - LL_2) = 91 > \chi^2_{2,99} = 9.21$.

**Table II.5**
**Multinomial Logit Estimates**

| Variable | Alternative 1 (1 accident) Original Coefficient | (t) | Transformed Coefficient | Alternative 2 (2 accidents or more) Original Coefficient | (t) | Transformed Coefficient |
|---|---|---|---|---|---|---|
| D | 0.0959 | (8.67)** | | 0.1337 | (5.35)** | |
| 0-1 | | | 0.0042 | | | 0.0002 |
| 1-5 | | | 0.0208 | | | 0.0013 |
| 5-31 | | | 0.3777 | | | 0.0518 |
| S | 0.5355 | (2.09)** | | 0.7073 | (1.38) | |
| 0-1 | | | 0.0307 | | | 0.0018 |
| 1-2 | | | 0.0478 | | | 0.0033 |
| 2-3 | | | 0.0702 | | | 0.0059 |
| Z | 0.3830 | (7.64)** | | 0.4780 | (3.75)** | |
| 0-1 | | | 0.0194 | | | 0.0010 |
| 1-2 | | | 0.0270 | | | 0.0015 |
| 2-6 | | | 0.2197 | | | 0.0170 |
| CONSTANT | -4.6116 | (-9.02)** | NC | -7.1174 | (-5.01)** | NC |
| A16- | -0.4123 | (-0.54) | -0.0155 | -5.8179 | (-0.25) | -0.0019 |
| A2024 | 0.0440 | (0.27) | 0.0021 | -0.8866 | (-1.85)* | -0.0013 |
| A2534 | -0.2931 | (-1.63) | -0.0124 | -1.7629 | (-3.06)** | -0.0025 |
| A3544 | -0.4756 | (-2.43)** | -0.0190 | -1.9913 | (-2.99)** | -0.0024 |
| A4554 | -0.4482 | (-2.18)** | -0.0175 | -2.3876 | (-3.22)** | -0.0024 |
| A5564 | -0.7113 | (-3.21)** | -0.0251 | -2.4728 | (-3.04)** | -0.0022 |
| A65+ | -0.6944 | (-2.68)** | -0.0240 | -2.3354 | (-2.52)** | -0.0019 |
| SEXM | 0.7790 | (9.30)** | 0.0335 | 1.2107 | (3.24)** | 0.0021 |
| EXP0 | 0.5150 | (1.11) | 0.0290 | 0.6040 | (0.48) | 0.0014 |
| EXPM1 | 0.2392 | (0.70) | 0.0120 | 0.0382 | (0.03) | 0.0001 |
| EXP2 | -0.4384 | (-1.63) | -0.0167 | 0.9395 | (0.98) | 0.0028 |
| EXP35 | -0.5978 | (-2.15)** | -0.0223 | 0.7716 | (0.77) | 0.0021 |
| EXP610 | -0.6286 | (-2.09)** | -0.0246 | 0.9485 | (0.88) | 0.0025 |
| EXP11 | -0.4826 | (-1.57) | -0.0220 | 0.8309 | (0.73) | 0.0016 |
| REG1 | 0.1747 | (0.99) | 0.0084 | 0.4144 | (0.75) | 0.0009 |
| REG2 | 0.0678 | (0.42) | 0.0031 | 0.1289 | (0.23) | 0.0003 |
| REG3 | 0.1456 | (1.46) | 0.0069 | 0.0759 | (0.21) | 0.0001 |
| REG4 | 0.2302 | (1.83)* | 0.0114 | -0.1531 | (-0.31) | -0.0003 |
| REG5 | 0.0255 | (0.14) | 0.0012 | -0.2407 | (-0.38) | -0.0004 |
| REG7 | 0.1195 | (1.01) | 0.0056 | 0.4175 | (1.16) | 0.0009 |
| REG8 | 0.0212 | (0.22) | 0.0010 | -0.2610 | (-0.72) | -0.0005 |
| REG9 | 0.4696 | (3.22)** | 0.0256 | 0.9776 | (2.32)** | 0.0028 |
| REG10 | 0.2334 | (1.27) | 0.0117 | 0.0992 | (0.16) | 0.0002 |
| REG11 | 0.3246 | (1.46) | 0.0170 | -0.3065 | (-0.30) | -0.0005 |

**Table II.5 (cont'd)**

| Variable | Alternative 1 (1 accident) Original Coefficient | (t) | Transformed Coefficient | Alternative 2 (2 accidents or more) Original Coefficient | (t) | Transformed Coefficient |
|---|---|---|---|---|---|---|
| RTSA | -0.5839 | (-2.61)** | -0.0239 | 0.7302 | (0.99) | 0.0017 |
| RTSB | 0.3707 | (0.68) | 0.0199 | -4.7744 | (-0.22) | -0.0019 |
| RTSGC | -0.3743 | (-1.15) | -0.0144 | 0.3971 | (0.34) | 0.0009 |
| RTSD | -0.1300 | (-0.47) | -0.0064 | 2.5122 | (3.64)** | 0.0196 |
| RTSH | 0.2431 | (1.03) | 0.0121 | 0.3783 | (0.43) | 0.0008 |
| RTSJ | 1.3778 | (1.76)* | 0.1173 | -2.3203 | (-0.04) | -0.0017 |
| RTSK | -0.9577 | (-0.95) | -0.0287 | -1.1703 | (-0.03) | -0.0013 |
| RTSM | -0.4174 | (-1.12) | -0.0159 | -7.1072 | (-0.31) | -0.0021 |
| RTSO | 0.3185 | (0.70) | 0.0166 | -0.1548 | (-0.13) | -0.0003 |
| RTSQ | -1.0159 | (-0.79) | -0.0297 | -1.2722 | (-0.02) | -0.0013 |
| RTSU | 0.7210 | (1.84)* | 0.0450 | 1.0183 | (0.87) | 0.0030 |
| RTSY | -2.0699 | (-2.04)** | -0.0415 | -5.0153 | (-0.22) | -0.0019 |
| RTSØ | -0.4383 | (-1.91)* | -0.0214 | 1.0706 | (1.37) | 0.0018 |
| CL1112 | 0.1200 | (0.52) | 0.0055 | 1.3103 | (2.82)** | 0.0049 |
| CL13 | -0.3465 | (-0.33) | -0.0134 | -4.1667 | (-0.19) | -0.0019 |
| CL21 | 0.2783 | (1.98)** | 0.0135 | 0.4812 | (1.11) | 0.0010 |
| CL22 | 0.9972 | (2.71)** | 0.0708 | 1.1778 | (1.34) | 0.0037 |
| CL31 | 0.6212 | (2.52)** | 0.0370 | 0.8989 | (1.56) | 0.0025 |
| CL42 | 0.1125 | (0.93) | 0.0050 | -0.1714 | (-0.47) | -0.0003 |
| CL54 | 0.0855 | (0.53) | 0.0038 | -0.1999 | (-0.43) | -0.0004 |
| CL55 | -1.6369 | (-1.57) | -0.0379 | -0.1010 | (-0.08) | -0.0001 |
| CL56 | 0.0262 | (0.02) | 0.0013 | -3.9273 | (-0.19) | -0.0018 |
| VALA | 0.9389E-3 | (0.81) | NC | -0.2512E-4 | (-0.01) | NC |
| VALB | 0.6760E-4 | (0.07) | NC | -0.2900E-3 | (-0.11) | NC |
| VALC | 4.7879E-3 | (5.63)** | NC | 0.8391E-3 | (0.46) | NC |

| | |
|---|---|
| LL at zero | = -4876.4237 |
| LL at convergence | = -4488.4113 |
| Log likelihood ratio | = 776.0248 |
| Number of observations | = 19,013 |
| Number of alternatives | = 3 |
| Number of parameters | = 106 |

* significant at 90%
** significant at 95%

**Table IL6**

**Maximum Likelihood Estimates of the Poisson Model with Individual Characteristics**

| Variable | Original Coefficient | (t) | Transformed Coefficients Prob1 | Prob2+ |
|---|---|---|---|---|
| D | 0.0832 | (10.23)** | | |
| 0-1 | | | 0.0039 | 0.0002 |
| 1-5 | | | 0.0186 | 0.0013 |
| 5-31 | | | 0.2702 | 0.1368 |
| S | 0.4624 | (2.49)** | | |
| 0-1 | | | 0.0271 | 0.0020 |
| 1-2 | | | 0.0396 | 0.0048 |
| 2-3 | | | 0.0547 | 0.0114 |
| Z | 0.3370 | (8.96)** | | |
| 0-1 | | | 0.0178 | 0.0011 |
| 1-2 | | | 0.0237 | 0.0022 |
| 2-6 | | | 0.1704 | 0.0516 |
| CONSTANT | -4.3318 | (-10.20)** | NC | NC |
| A16- | -0.6992 | (-0.96) | -0.0247 | -0.0010 |
| A2024 | -0.0959 | (-0.69) | -0.0044 | -0.0002 |
| A2534 | -0.4738 | (-3.06)** | -0.0208 | -0.0010 |
| A3544 | -0.6427 | (-3.76)** | -0.0262 | -0.0012 |
| A4554 | -0.6696 | (-3.71)** | -0.0261 | -0.0012 |
| A5564 | -0.8841 | (-4.51)** | -0.0314 | -0.0013 |
| A65+ | -0.8542 | (-3.71)** | -0.0296 | -0.0012 |
| SEXM | 0.7810 | (10.13)** | 0.0356 | 0.0020 |
| EXP0 | 0.5620 | (1.43) | 0.0341 | 0.0027 |
| EXPM1 | 0.2531 | (0.85) | 0.0134 | 0.0008 |
| EXP2 | -0.1892 | (-0.81) | -0.0084 | -0.0004 |
| EXP35 | -0.3806 | (-1.56) | -0.0161 | -0.0008 |
| EXP610 | -0.3450 | (-1.31) | -0.0152 | -0.0008 |
| EXP11 | -0.2613 | (-0.96) | -0.0125 | -0.0007 |
| REG1 | 0.1716 | (1.12) | 0.0088 | 0.0005 |
| REG2 | 0.0828 | (0.59) | 0.0041 | 0.0002 |
| REG3 | 0.1033 | (1.17) | 0.0051 | 0.0003 |
| REG4 | 0.1439 | (1.28) | 0.0073 | 0.0004 |
| REG5 | -0.0294 | (-0.19) | -0.0014 | -0.0001 |
| REG7 | 0.1409 | (1.40) | 0.0071 | 0.0004 |
| REG8 | -0.0238 | (-0.28) | -0.0011 | -0.0001 |
| REG9 | 0.4677 | (3.78)** | 0.0271 | 0.0020 |
| REG10 | 0.1750 | (1.09) | 0.0090 | 0.0005 |
| REG11 | 0.2147 | (1.07) | 0.0113 | 0.0007 |

**Table II.6 (cont'd)**

| Variable | Original Coefficient | (t) | Transformed Coefficients Prob1 | Prob2+ |
|---|---|---|---|---|
| RTSA | -0.3324 | (-1.72)* | -0.0150 | -0.0008 |
| RTSB | 0.2342 | (0.45) | 0.0124 | 0.0008 |
| RTSGC | -0.1850 | (-0.64) | -0.0082 | -0.0004 |
| RTSD | 0.2941 | (1.28) | 0.0159 | 0.0010 |
| RTSH | 0.2218 | (1.07) | 0.0116 | 0.0007 |
| RTSJ | 1.4493 | (2.04)** | 0.1299 | 0.0206 |
| RTSK | -0.9161 | (-0.98) | -0.0296 | -0.0011 |
| RTSM | -0.5860 | (-1.65)* | -0.0219 | -0.0009 |
| RTSO | 0.2402 | (0.63) | 0.0128 | 0.0008 |
| RTSQ | -1.2459 | (-1.06) | -0.0354 | -0.0012 |
| RTSU | 0.4873 | (1.50) | 0.0289 | 0.0022 |
| RTSY | -1.9646 | (-1.95)* | -0.0434 | -0.0013 |
| RTSØ | -0.1831 | (-0.92) | -0.0090 | -0.0005 |
| CL1112 | 0.4293 | (2.56)** | 0.0248 | 0.0018 |
| CL13 | -0.4499 | (-0.45) | -0.0177 | -0.0008 |
| CL21 | 0.2745 | (2.29)** | 0.0141 | 0.0008 |
| CL22 | 0.8311 | (2.98)** | 0.0575 | 0.0055 |
| CL31 | 0.5504 | (2.98)** | 0.0336 | 0.0026 |
| CL42 | 0.0681 | (0.66) | 0.0032 | 0.0002 |
| CL54 | 0.0230 | (0.17) | 0.0011 | 0.0001 |
| CL55 | -0.6993 | (-1.18) | -0.0247 | -0.0010 |
| CL56 | -0.2090 | (-0.21) | -0.0091 | -0.0004 |
| VALA | 0.3382 | (0.90) | NC | NC |
| VALB | 0.0660 | (0.22) | NC | NC |
| VALC | 1.2803 | (5.27)** | NC | NC |
| Average individual | | | 0.0504 | 0.0014 |
| LL | -4563.86 | | | |
| Number of observations | 19,013 | | | |
| Number of parameters | 53 | | | |

* = significant at 90%; ** = significant at 95%

**Table II.7**

**Maximum Likelihood Estimates of the Negative Binomial Model with Individual Characteristics**

| Variable | Original Coefficient | (t) | Transformed Coefficients Prob1 | Prob2+ |
|---|---|---|---|---|
| D | 0.0863 | (9.66)** | | |
| 0-1 | | | 0.0038 | 0.0003 |
| 1-5 | | | 0.0183 | 0.0019 |
| 5-31 | | | 0.2219 | 0.1671 |
| S | 0.4800 | (2.19)** | | |
| 0-1 | | | 0.0266 | 0.0029 |
| 1-2 | | | 0.0380 | 0.0070 |
| 2-3 | | | 0.0501 | 0.0160 |
| Z | 0.3488 | (8.40)** | | |
| 0-1 | | | 0.0175 | 0.0016 |
| 1-2 | | | 0.0230 | 0.0032 |
| 2-6 | | | 0.1486 | 0.0684 |
| CONSTANT | -4.3138 | (-8.43)** | NC | NC |
| A16- | -0.7091 | (-0.92) | -0.0241 | -0.0015 |
| A2024 | -0.1048 | (-0.70) | -0.0046 | -0.0003 |
| A2534 | -0.4874 | (-2.89)** | -0.0203 | -0.0016 |
| A3544 | -0.6581 | (-3.47)** | -0.0256 | -0.0018 |
| A4554 | -0.6801 | (-3.43)** | -0.0253 | -0.0017 |
| A5564 | -0.8949 | (-4.25)** | -0.0304 | -0.0019 |
| A65+ | -0.8522 | (-3.66)** | -0.0284 | -0.0017 |
| SEXM | 0.7768 | (9.66)** | 0.0337 | 0.0027 |
| EXP0 | 0.5493 | (1.20) | 0.0311 | 0.0034 |
| EXPM1 | 0.2573 | (0.71) | 0.0129 | 0.0012 |
| EXP2 | -0.1854 | (-0.65) | -0.0079 | -0.0006 |
| EXP35 | -0.3850 | (-1.32) | -0.0155 | -0.0011 |
| EXP610 | -0.3423 | (-1.10) | -0.0144 | -0.0011 |
| EXP11 | -0.2565 | (-0.80) | -0.0117 | -0.0009 |
| REG1 | 0.1791 | (1.13) | 0.0087 | 0.0007 |
| REG2 | 0.0796 | (0.56) | 0.0037 | 0.0003 |
| REG3 | 0.1059 | (1.14) | 0.0050 | 0.0004 |
| REG4 | 0.1451 | (1.20) | 0.0069 | 0.0006 |
| REG5 | -0.0251 | (-0.15) | -0.0011 | -0.0001 |
| REG7 | 0.1408 | (1.36) | 0.0067 | 0.0006 |
| REG8 | -0.0236 | (-0.26) | -0.0011 | -0.0001 |
| REG9 | 0.4631 | (3.50)** | 0.0252 | 0.0027 |
| REG10 | 0.1836 | (1.09) | 0.0090 | 0.0008 |
| REG11 | 0.2188 | (1.02) | 0.0109 | 0.0010 |

**Table II.7 (cont'd)**

| Variable | Original Coefficient | (t) | Transformed Coefficients Prob1 | Prob2+ |
|---|---|---|---|---|
| RTSA | -0.3450 | (-1.59) | -0.0148 | -0.0012 |
| RTSB | 0.2412 | (0.45) | 0.0121 | 0.0012 |
| RTSGC | -0.1871 | (-0.56) | -0.0079 | -0.0006 |
| RTSD | 0.2795 | (1.22) | 0.0142 | 0.0014 |
| RTSH | 0.2189 | (0.85) | 0.0108 | 0.0010 |
| RTSJ | 1.4465 | (1.49) | 0.1154 | 0.0262 |
| RTSK | -0.9376 | (-0.74) | -0.0290 | -0.0017 |
| RTSM | -0.5990 | (-1.47) | -0.0214 | -0.0014 |
| RTSO | 0.2458 | (0.63) | 0.0124 | 0.0012 |
| RTSQ | -1.2141 | (-0.84) | -0.0338 | -0.0018 |
| RTSU | 0.5083 | (1.44) | 0.0285 | 0.0031 |
| RTSY | -1.9666 | (-1.84)* | -0.0421 | -0.0019 |
| RTSØ | -0.1936 | (-0.87) | -0.0091 | -0.0007 |
| CL1112 | 0.4270 | (2.54)** | 0.0231 | 0.0025 |
| CL13 | -0.4715 | (-0.42) | -0.0176 | -0.0011 |
| CL21 | 0.2570 | (1.99)* | 0.0124 | 0.0011 |
| CL22 | 0.8430 | (2.64)** | 0.0542 | 0.0075 |
| CL31 | 0.5449 | (2.64)** | 0.0310 | 0.0035 |
| CL42 | 0.0587 | (0.53) | 0.0026 | 0.0002 |
| CL54 | 0.0333 | (0.24) | 0.0015 | 0.0001 |
| CL55 | -0.7309 | (-1.11) | -0.0246 | -0.0015 |
| CL56 | -0.2455 | (-0.18) | -0.0101 | -0.0007 |
| VALA | 0.3624 | (0.76) | NC | NC |
| VALB | 0.0131 | (0.04) | NC | NC |
| VALC | 1.3066 | (4.78)** | NC | NC |
| a | 0.4609 | (3.18)** | | |
| Average individual | | | 0.0492 | 0.0020 |
| LL | | -4556.43 | | |
| Number of observations | | 19,013 | | |
| Number of parameters | | 54 | | |

* = significant at 90%; ** = significant at 95%

**Table II.8**

**Quasi-Generalized Pseudo Maximum Likelihood Estimates**

| Variable | Coefficient | (t) |
|---|---|---|
| D | 0.0839 | (10.64)** |
| S | 0.4647 | (2.32)** |
| Z | 0.3400 | (9.35)** |
| CONSTANT | -4.3260 | (-8.82)** |
| A16- | -0.7017 | (-0.93) |
| A2024 | -0.0981 | (-0.70) |
| A2534 | -0.4777 | (-2.99)** |
| A3544 | -0.6470 | (-3.58)** |
| A4554 | -0.6728 | (-3.56)** |
| A5564 | -0.8873 | (-4.39)** |
| A65+ | -0.8550 | (-3.80)** |
| SEXM | 0.7799 | (9.96)** |
| EXP0 | 0.5580 | (1.28) |
| EXPM1 | 0.2536 | (0.73) |
| EXP2 | -0.1896 | (-0.71) |
| EXP35 | -0.3823 | (-1.38) |
| EXP610 | -0.3442 | (-1.17) |
| EXP11 | -0.2601 | (-0.86) |
| REG1 | 0.1727 | (1.14) |
| REG2 | 0.0819 | (0.60) |
| REG3 | 0.1036 | (1.16) |
| REG4 | 0.1438 | (1.25) |
| REG5 | -0.0289 | (-0.18) |
| REG7 | 0.1406 | (1.43) |
| REG8 | -0.0242 | (-0.28) |
| REG9 | 0.4665 | (3.73)** |
| REG10 | 0.1762 | (1.10) |
| REG11 | 0.2152 | (1.05) |
| RTSA | -0.3346 | (-1.62) |
| RTSB | 0.2356 | (0.45) |
| RTSGC | -0.1852 | (-0.58) |
| RTSD | 0.2912 | (1.32) |
| RTSH | 0.2214 | (0.90) |
| RTSJ | 1.4474 | (1.60) |
| RTSK | -0.9169 | (-0.78) |
| RTSM | -0.5881 | (-1.48) |
| RTSO | 0.2405 | (0.64) |
| RTSQ | -1.2402 | (-0.92) |
| RTSU | 0.4922 | (1.53) |
| RTSY | -1.9652 | (-1.86)* |
| RTSØ | -0.1848 | (-0.87) |

**Table IL8 (cont'd)**

| Variable | Coefficient | (t) |
|---|---|---|
| CL1112 | 0.4284 | (2.79)** |
| CL13 | -0.4517 | (-0.41) |
| CL21 | 0.2699 | (2.22)** |
| CL22 | 0.8332 | (2.87)** |
| CL31 | 0.5480 | (2.95)** |
| CL42 | 0.0653 | (0.63) |
| CL54 | 0.0254 | (0.19) |
| CL55 | -0.7042 | (-1.15) |
| CL56 | -0.2110 | (-0.17) |
| VALA | 0.3441 | (0.76) |
| VALB | 0.0537 | (0.17) |
| VALC | 1.2855 | (4.98)** |

| | |
|---|---|
| Objective function: | -4492.99 |
| Nb of parameters: | 53 |
| Nb of observations: | 19 013 |

Table II.9

The Otherwise Average Individual's Probability of Accident as a Function of Cumulated Demerit Points D

| Demerit Points D During 1980-1981 and 1981-1982 | Probability of at Least 1 Accident in 1982-1983 (for S = Z = 0) | | | |
|---|---|---|---|---|
| | PROBIT | POISSON | NEGBIN | MLOGIT |
| 0 | 0.043 | 0.046 | 0.045 | 0.043 |
| 5 | 0.075 | 0.068 | 0.068 | 0.068 |
| 10 | 0.123 | 0.102 | 0.101 | 0.105 |
| 15 | 0.188 | 0.150 | 0.150 | 0.161 |
| 20 | 0.271 | 0.218 | 0.218 | 0.239 |

Table II.10

The Otherwise Average Individual's Probability of Accident as a Function of License Suspensions S

| Number of Suspensions S During 1981-1982 | Probability of at Least 1 Accident in 1982-1983 (for D = Z = 0) | | | |
|---|---|---|---|---|
| | PROBIT | POISSON | NEGBIN | MLOGIT |
| 0 | 0.043 | 0.046 | 0.045 | 0.043 |
| 1(VALB = 281)[1] | 0.079 | 0.071 | 0.071 | 0.071 |
| 2(VALB = 197)[1] | 0.134 | 0.108 | 0.110 | 0.114 |

[1] The driver cannot drive during a suspension.

Table II.11

The Otherwise Average Individual's Probability of Accident as a Function of Past Involvements in Accidents Z

| Number of Accidents Z During 1980–1981 and 1981–1982 | Probability of at Least 1 Accident in 1982–1983 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | For D=S=0 | | | | For D=5 and S=1 (VALB=281) | | | |
| | Probit | Poisson | Negbin | Mlogit | Probit | Poisson | Negbin | Mlogit |
| 0 | 0.043 | 0.0455 | 0.0447 | 0.0427 | 0.128 | 0.1050 | 0.1058 | 0.1100 |
| 1 | 0.066 | 0.0632 | 0.0625 | 0.0616 | 0.177 | 0.1440 | 0.1452 | 0.1542 |
| 2 | 0.098 | 0.0873 | 0.0869 | 0.0882 | 0.238 | 0.1957 | 0.1968 | 0.2119 |
| 3 | 0.140 | 0.1202 | 0.1200 | 0.1248 | 0.307 | 0.2629 | 0.2625 | 0.2842 |
| 4 | 0.192 | 0.1642 | 0.1639 | 0.1737 | 0.385 | 0.3477 | 0.3429 | 0.3696 |
| 5 | 0.255 | 0.2221 | 0.2209 | 0.2366 | 0.468 | 0.4503 | 0.4366 | 0.4643 |

Table II.12

The Otherwise Average Individual's Expected Number of Accidents E(ACC) as a Function of Cumulated Demerit Points D

| Demerit Points D During 1980-1981 and 1981-1982 | E(ACC) in 1982-1983 (for S = Z = 0) | | |
|---|---|---|---|
| | POISSON | NEGBIN | MLOGIT |
| 0 | 0.047 | 0.046 | 0.044 |
| 5 | 0.071 | 0.071 | 0.071 |
| 10 | 0.107 | 0.110 | 0.111 |
| 15 | 0.162 | 0.169 | 0.171 |
| 20 | 0.246 | 0.260 | 0.256 |

Table II.13

The Otherwise Average Individual's Expected Number of Accidents E(ACC) as a Function of License Suspensions S

| Number of Suspensions S During 1981-1982 | E(ACC) in 1982-1983 (for D = Z = 0) | | |
|---|---|---|---|
| | POISSON | NEGBIN | MLOGIT |
| 0 | 0.047 | 0.046 | 0.044 |
| 1(VALB = 281)[1] | 0.073 | 0.075 | 0.074 |
| 2(VALB = 197)[1] | 0.115 | 0.120 | 0.120 |

[1] The driver cannot drive during a suspension.

Table II.14

**The Otherwise Average Individual's Expected Number of Accidents E(ACC) as a Function of Past Involvements in Accidents Z**

| Number of Accidents Z During 1980–1981 and 1981–1982 | E(ACC) in 1982–1983 | | | | | |
|---|---|---|---|---|---|---|
| | For D=S=0 | | | For D=5 and S=1 (VALB=281) | | |
| | Poisson | Negbin | Mlogit | Poisson | Negbin | Mlogit |
| 0 | 0.0466 | 0.0462 | 0.0443 | 0.1110 | 0.1148 | 0.1157 |
| 1 | 0.0652 | 0.0655 | 0.0641 | 0.1554 | 0.1627 | 0.1629 |
| 2 | 0.0914 | 0.0929 | 0.0921 | 0.2177 | 0.2306 | 0.2251 |
| 3 | 0.1280 | 0.1317 | 0.1308 | 0.3050 | 0.3269 | 0.3035 |
| 4 | 0.1793 | 0.1866 | 0.1828 | 0.4272 | 0.4634 | 0.3971 |
| 5 | 0.2512 | 0.2645 | 0.2503 | 0.5984 | 0.6568 | 0.5019 |